



Cell

Volume 163
Number 2

October 8, 2015

www.cell.com



Balancing Balance

BASALDELLA ET AL., PAGE 301

The vestibular and proprioceptive systems control body balance by targeting motor neuron populations regulating functionally distinct muscles and fiber types, and their inputs are shaped by competitive interactions between these two systems.

Sniffing Out Specificity

FU ET AL., PAGE 313

A technique for linking specific chemical compounds within mixtures from bodily fluids to functionally distinct classes of sensory neurons sets the stage for assigning discrete functional identities to individual vomeronasal neurons as shown for a new family of murine female pheromones.

Interferon with Neurodegeneration

EJLERSKOV ET AL., PAGE 324

Lack of the immunomodulatory cytokine interferon- β causes spontaneous neurodegeneration resembling sporadic Lewy body and Parkinson's disease dementia due to defects in neuronal autophagy.

Fat Chance for Survival

BAILEY ET AL., PAGE 340

Lipid droplets can act as anti-oxidant organelles to protect *Drosophila* neural stem cells from hypoxia-triggered ROS by sequestering damage-vulnerable polyunsaturated fatty acids from plasma membranes.

Slithering Builds an Organ

KUO ET AL., PAGE 394

A form of migration named "slithering," where cells undergo transient EMT, allows rapid assembly of pulmonary neurosensory organs and may explain why lung cancers arising from neuroendocrine cells are highly metastatic.

Purposeful Protein Aggregation

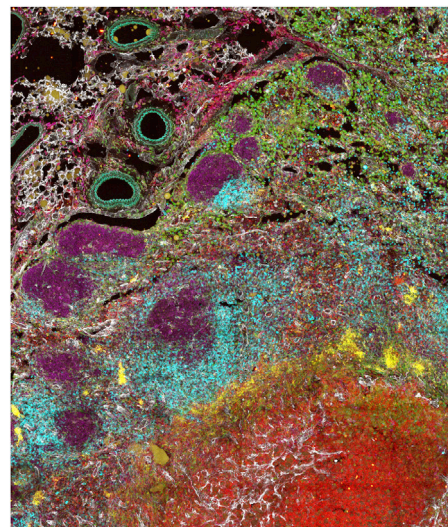
BERCHOWITZ ET AL., PAGE 406

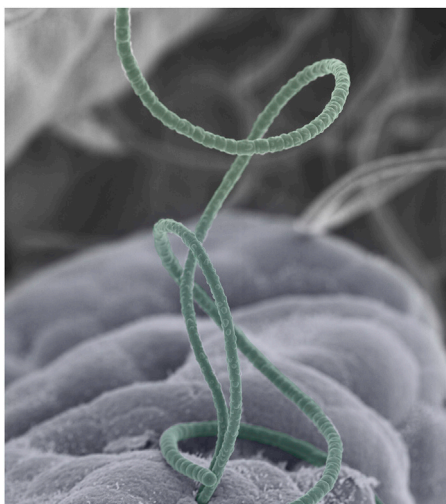
Amyloid-like aggregation of the RNA-binding protein Rim4 controls gametogenesis by repressing message-specific translation.

A Lasting Scar of Acute Infection

DA FONSECA ET AL., PAGE 354

Following a single episode of acute gastrointestinal infection, microbiota-dependent inflammation of fat tissue and damage to the lymphatic system disrupt the communication between the immune system and the tissue, causing an "immunological scar" that compromises tissue immunity and homeostasis in the long term.





Stick to Your Gut

SANO ET AL., PAGE 381

ATARASHI ET AL., PAGE 367

The adhesion of specific members of the gut microbiome to epithelial cells in the small intestine initiates a signaling axis involving CD4⁺ T cells and serum amyloid A to generate mucosally protective Th17 cells, thereby highlighting the importance of location in the regulation of gut homeostasis.

Protease Lego

JOSHI ET AL., PAGE 419

Sequential protein degradation driving the *Caulobacter* cell cycle relies on an adaptor hierarchy—priming of the protease by one adaptor recruits additional adaptors, and substrate degradation occurs based on the degree of adaptor assembly.

Snapping into Shape

TOMKO ET AL., PAGE 432

The base and lid complexes of the functional proteasome can only associate once a single alpha helix from the last-to-be incorporated subunit of the lid complex snaps into place and triggers a large-scale conformational change.

siRNAs One By One

ZHAI ET AL., PAGE 445

Rather than a single RNA transcript giving rise to multiple siRNAs, Pol IV transcribes multiple short RNAs that are further processed to generate mature products, thereby demonstrating a “one precursor, one product” model for Pol IV-dependent siRNA biogenesis in *Arabidopsis*.

A Brainy Network

MARKRAM ET AL., PAGE 456

A large-scale digital reconstruction and simulation of neocortical microcircuitry suggests cellular and synaptic mechanisms that dynamically reconfigure the state of the network to support diverse strategies of information processing.

Stratifying Breast Cancer

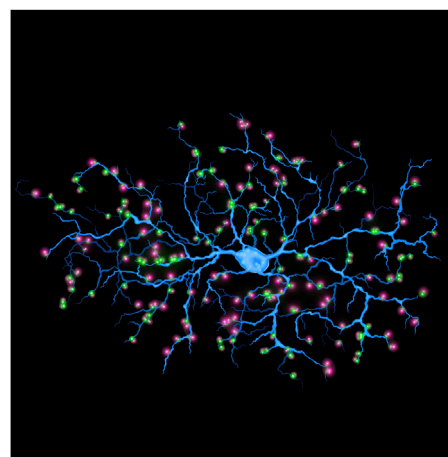
CIRIELLO ET AL., PAGE 506

A comprehensive analysis of breast tumor samples determines invasive lobular carcinoma as a molecularly distinct disease with characteristic genetic features, providing key information for patient stratification that may allow a more informed clinical follow-up.

Nanoscale Neurocircuitry in 3D

SIGAL ET AL., PAGE 493

Super-resolution imaging with multi-color reconstruction of synapses in brain tissue enables interrogation of neural circuitry at nanoscale.



A Spotless Mind

Memories are difficult to pin down, in part because there is no single kind—they can be fragmentary or interwoven, stable or ephemeral, easily recalled or long dormant. Exploring how learning and memory are encoded by changes in neuronal structure is a pursuit that supplants the abstract with something tangible, a toehold for understanding one of the most fundamental of experiences.

Seeking out where a specific memory is made and systematically erasing it by manipulating ensembles of synaptic structures is the great achievement of recent work by [Hayashi-Takagi et al. \(2015\)](#). They cleverly create a probe that contains a photoactivatable form of the small GTPase Rac1 (which when active induces spine shrinkage) and link it to a deletion mutant of PSD-95, which, together with a dendritic targeting element (DTE), endows the probe with the property of accumulating at activated synapses. With this combo of features, the probe targets recently potentiated spines and upon photostimulation will induce spine shrinkage. Using it, the authors are able to see the extensive synaptic remodeling that occurs in the motor cortex as mice learn a task involving balance and motor coordination. When the probe is activated following training, the mastery gained by the mice is lost, whereas targeting spines activated by a different motor learning task does not impact learning of the original one, despite the spatial overlap of the assemblies of activated synapses.

The authors estimate that the erasure of the acquired motor memory entails removal of ~410,000 spines from ~4,700 neurons, providing a glimpse of the complexity of neuronal structural changes responsible for task-specific learning. This raises interesting questions. How are these neurons linked in circuits? And how equivalent are the contributions of individual spines to the acquired memory? Ultimately, it will also be interesting to learn how this scale of structural change compares to that which might be formed in other brain regions by exposure to objects and events. Might it be possible to accelerate learning with photoactivatable probes designed to augment the formation of these synaptic ensembles? Or perhaps to engineer proteins that won't require photostimulation for activation, but might instead be triggered remotely? Studies have achieved remote control of channel proteins by radio waves using genetically encoded nanoparticles ([Stanley et al., 2012](#); [Stanley et al., 2015](#)). Such an approach, if possible, could target neurons more spatially separated than the ensembles studied by Hayashi-Takagi et al. Might it even be possible to make mice with a reset switch for specific episodic memories?

Controlling behavior in mice with optogenetic tools is being deployed with ever greater precision and can now be used to intervene in behaviors once thought experimentally intractable, a recent example being the reactivation of neurons linked to a positive experience as a means of alleviating depression-like symptoms ([Ramirez et al., 2015](#)). Such interventions, however, are not yet possible in humans or primates, though steps in this direction are being made, including a recent study by [Inoue et al. \(2015\)](#) exploring optogenetic stimulation in macaque monkeys. Inoue et al. introduced channelrhodopsin via an adeno-associated virus



Visualizing neuronal spine potentiation. Image courtesy of A. Hayashi-Takagi.

vector to the frontal eye field, a region of the primate frontal cortex that is part of the oculomotor system. The optogenetic stimulation of channelrhodopsin-expressing axons that project from this area to the superior colliculus triggers rapid involuntary eye movements known as saccades, demonstrating the feasibility of pathway-selective optogenetic stimulation in primates. Casting our glance to the scientific horizon, it seems that such studies bode well for translating the panoply of tools used in rodents to studies of behavior in a wider range of models, particularly those amenable to genetic modification (including macaques, [Niu et al., 2014](#)) or one day even to therapeutic applications in humans.

REFERENCES

- Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y.I., Loshbaugh, A.L., Kuhlman, B., Hahn, K.M., and Kasai, H. (2015). *Nature* 525, 333–338.
- Inoue, K., Takada, M., and Matsumoto, M. (2015). *Nat. Commun.* 6, 8378.
- Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., et al. (2014). *Cell* 156, 836–843.
- Ramirez, S., Liu, X., MacDonald, C.J., Moffa, A., Zhou, J., Redondo, R.L., and Tonegawa, S. (2015). *Nature* 522, 335–339.
- Stanley, S.A., Gagner, J.E., Damanpour, S., Yoshida, M., Dordick, J.S., and Friedman, J.M. (2012). *Science* 336, 604–608.
- Stanley, S.A., Sauer, J., Kane, R.S., Dordick, J.S., and Friedman, J.M. (2015). *Nat. Med.* 21, 92–98.

Robert P. Kruger

Pied Piper of Neuroscience

Principles of Neurobiology

Author: Liqun Luo

New York, NY, USA: Garland Science (2015).

645 pp. \$100.00

Forgive us for not disclosing the year, but when one of us was excited about the brain at a young age (and the other just born), neuroscience as an integrated discipline was still so new that one could hardly find a textbook.

Neuroscience was then scattered in the standard medical curriculum of anatomy, embryology, histology, biochemistry, physiology, pharmacology, pathology, neurology, and psychiatry. Some teachers literally extracted relevant materials from those traditional disciplines before adding a prefix “neuro-” to most of them. The exception was Harvard University, where Steve Kuffler and John Nicholls published *From Neuron to Brain* in 1976, basically using their teaching notes. For the progeny of Harvard who worshiped Kuffler, *From Neuron to Brain* was automatically embraced, though its style was highly personal and its content highly selective. Kuffler and Nicholls imprinted what they liked onto their students, who later spread the book when they established laboratories or departments of neurobiology outside Harvard.

When Eric Kandel and James Schwartz published *Principles of Neural Science* in 1981, it solved the problem of comprehensiveness and became the standard neuroscience textbook. In its third edition in 1991, Tom Jessell strengthened the coverage on neural development, which, at that time, was taught mainly from *Principles of Neural Development* by Dale Purves and Jeff Lichtman (1985). More textbooks, including *Fundamental Neuroscience* by Squire et al., *Neuroscience* by Purves et al., and *Neuroscience Exploring the Brain* by Bear et al., followed and each has been used by a significant readership.

Why should anyone read Dr. Liqun Luo's *Principles of Neurobiology*? Isn't it suicidal to write a textbook in an already crowded market?

The new book is distinct in its philosophy and authorship, which makes it val-

uable for teaching and a pleasure to read. Luo shares with us the belief that students should not be taught just facts and knowledge but also how knowledge is obtained and how principles are derived. He advocates discovery-based teaching in his own course at Stanford and Peking Universities. This book is the product of his hard work to offer discovery-based teaching to more students. The target readership is senior undergraduate students interested in research, graduate students interested in neuroscience, and the minority of medical students who still intend to pursue research alongside clinical practice.

It is an immense task for any individual to write a textbook for a discipline so large now that each recent annual meeting for the Society for Neuroscience in the United States alone has more than 30,000 attendees. If one wants to be sane and preserve one's own time for research, one should ask others to join in writing a book. These partially explain why all major neuroscience textbooks are multi-authored. Multiple authors also ensure expertise in diverse fields. Nevertheless, the

advantage of single authorship is also obvious, as long as the author has the intellectual breadth and depth to cover different areas. A consistent writing style and a coherent, logical flow will certainly be helpful to readers. Luo has been heroic in writing the book over 5 years as a single author. In 13 chapters and 645 pages, he provides a masterpiece highlighting principles of how the nervous system transforms and governs our sensation, action, memory, and thought.

Reading *Principles of Neurobiology* is like enjoying a beautiful piece of music. Luo often begins each section with interesting stories or classic experiments and ends with his comments on important but unsolved questions. We include a few stories here.

Chapter 1 showcases nature versus nurture: a young chick exhibits innate escape behavior if an object flies toward it resembling a hawk but not when the same object flies toward it resembling a goose; juvenile barn owls, after wearing artificial prisms, could adjust their auditory map to match an altered visual map. Luo uses these stories to entice students to explore neurobiological mechanisms of behavior: how experiences remodel and sculpt neural circuits formed by genetic instruction.

Chapters 2 to 11 form the core: signaling within neurons and across synapses, sensory/motor systems and how they wire together, sexual behavior, learning/memory and synaptic plasticity, and brain disorders. Unlike other neuroscience textbooks in which molecular, cellular, systems, and developmental neurobiology are treated separately, most topics in this book are question oriented with interdisciplinary approaches intertwined to address the same question. For example, in the section on olfaction, the author starts with an intriguing and remarkable story of salmon homing. A 1976 field study demonstrates that adult salmon rely on the imprinted memory of smell to guide their homing to streams even after they have migrated many miles away. How are odor stimuli transformed into electrical signals? How do these signals get processed, organized, and transmitted to the brain to form internal representation? How can some odors evoke stereotypic innate behaviors while others can be flexibly associated with learned

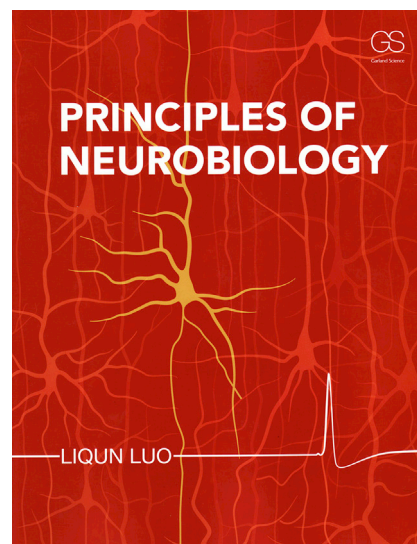




Image credit: Zhi Ye, Peking University.

behaviors? How do distinct neural maps form during development and then how are they shaped by experiences? The author logically takes readers through the journey of major discoveries, colorfully filled with chemistry, molecular biology, genetics, electrophysiology, optical imaging, and behavioral assays. Instead of simply remembering facts or conclusions, students can have fun deriving conclusions from the original data. The piper is getting the student to follow with action.

All sensory systems convert external stimuli into an internal representation in the brain, but they differ dramatically in the physical nature of the stimuli and the required processing. For example, vision has to convey color, as well as spatial information, while the auditory system requires frequency separation and temporal precision. Are there the general principles shared between olfactory, visual, auditory, taste, and somatosensory systems?

What are their unique mechanisms? Again, the discovery-based approach encourages the reader to solve the puzzles one by one and arrive at a general understanding of the sensory systems. This is but one of the many storylines that are thought provoking and enjoyable to read.

The book ends with “evolution of the nervous system” (Chapter 12) and “ways of exploring” (Chapter 13). The evolution section asks how ion channels or sensory GPCRs evolved and when and how the mammalian neocortex underwent rapid expansion. It places neurobiology in a perspective where diverse and independent strategies in different organisms were created to solve the common problem.

The last chapter centers on technologies in neurobiology. As Sydney Brenner remarked in 1980, “progress in science depends on new techniques, new discoveries, and new ideas, probably in that order.” In fact, the career of the author

himself partly echoes the comment. From the late 1990s to now, Luo has invented genetic methods in flies and mice to label and manipulate single neurons and to trace *trans*-synaptic connections, thus in an excellent position to provide an authoritative summary of neurotechniques, beyond those of optogenetics and CRISPR/Cas9. Regrettably, however, a bias is shown by the absence of biochemical approaches to neurobiology, which have played an important role in neuroscience such as the discoveries of the nerve growth factor and netrin.

Taken together, Luo's book is a piper's flute for neuroscience students. As a single author taking up the challenge of discovery-based teaching, Luo has included a judicious choice of materials and achieved a fine balance of breadth and depth. *Principle of Neurobiology* will help teachers to inspire curious students. Once the students find neurobiology irresistible, they will find that no one can cover neuroscience in a single textbook, and motivated students can go on his/her own path of learning more about neuroscience by reading other books that have upward of 1,700 pages and nearly 80 authors, specialized books such as *Cognitive Neuroscience* by Michael Gazzaniga, or better still original research papers.

Most scientists are afraid of writing textbooks, both because precious time can be spent on doing research and writing grants and because the reputation of authors of textbooks in the natural sciences almost always precedes, rather than follows, a textbook. Luo, through his *Principles of Neurobiology*, has set an example of how an outstanding scientist cares about education and writes a textbook to inspire students to pursue discoveries. We hope that students will follow this piper, happily and fruitfully.

Yulong Li^{1,2,3,*} and Yi Rao^{1,2,3,*}

¹State Key Laboratory of Membrane Biology, Peking University School of Life Sciences, Beijing 100871, China

²PKU-IDG/McGovern Institute for Brain Research, Beijing 100871, China

³Peking-Tsinghua Center for Life Sciences, Beijing 100871, China

*Correspondence: yulongli@pku.edu.cn

(Y.L.), yrao@pku.edu.cn (Y.R.)

<http://dx.doi.org/10.1016/j.cell.2015.09.039>

How Brain Fat Conquers Stress

Michael A. Welte^{1,*}

¹Department of Biology, University of Rochester, RC Box 270211, 317 Hutchison Hall, Rochester, NY 14627, USA

*Correspondence: michael.welte@rochester.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.046>

A new paper by Bailey et al. reveals that lipid droplets, crucial organelles for energy storage, can also protect against oxidative stress. In *Drosophila* larvae, lipid droplets in glia allow neuronal stem cells to keep proliferating under hypoxic conditions. Protection likely involves sequestering vulnerable membrane lipids away from reactive oxygen species.

Cells routinely encounter highly reactive oxygen derivatives, from endogenous and exogenous sources. These reactive oxygen species (ROS) mediate intracellular signaling but in excess cause severe damage to DNA, proteins, and lipids. Organisms combat such oxidative stress with a myriad of strategies, including small-molecule antioxidants and detoxifying enzymes. Somewhat paradoxically, increased ROS production also occurs in cells exposed to reduced oxygen levels (hypoxia), a particular problem for stem cells, as they tend to reside in hypoxic niches. A new paper in this issue of *Cell* (Bailey et al., 2015) uncovers a novel antioxidant strategy for stem cell protection. Here, glial cells harness fat storage organelles to protect neighboring neuronal stem cells, likely by sequestering vulnerable lipids away from membranes. This elegant paper sheds new light on lipid droplets, stem cells, and stress protection.

Lipid droplets are the sites where cells store fat (Walther and Farese, 2012). Here, a layer of phospholipids and proteins surrounds a core of neutral lipids (e.g., triglycerides). Lipid droplets play central roles in lipid metabolism, storing energy, providing precursors for membrane lipids, and sequestering toxic lipids. Not surprisingly, droplet dysfunction is linked to many human diseases, including fatty liver, diabetes, and cardiovascular disease.

Do lipid droplets also contribute to the complex lipid metabolism of the nervous system? So far, this possibility is poorly explored (Welte, 2015), though in adult flies and mice, inappropriate droplet accumulation in neurons or in glia is linked to neurodegeneration (Inloes et al., 2014; Liu et al., 2015). In the CNS of *Drosophila* larvae, droplets are abundant during

normal development, though only in certain types of glia (Kis et al., 2015). Bailey et al. (2015) discovered that oxidative stress, as a result of hypoxia or ROS-generating chemicals, vastly increases droplet number specifically in the CNS.

These larval glia serve as a niche for neuronal stem cells, the neuroblasts (Figure 1, left). Under low-nutrient conditions, they secrete signals that allow neuroblasts to keep proliferating, even as other cell lineages stop dividing (Cheng et al., 2011). Such “brain sparing” selectively protects the CNS, so that larvae can still develop into fully functional, albeit smaller, adults. By abolishing droplets specifically in glia, Bailey et al. (2015) now find that droplet accumulation promotes brain sparing under conditions of oxidative stress: glial lipid droplets allow neuroblasts to keep proliferating under hypoxic conditions.

How does this protection work? One important type of ROS-induced damage is lipid peroxidation. Using a fluorescent sensor, the authors found that lipid peroxidation is higher in membranes than inside droplets and that membrane damage goes up dramatically when glia lack lipid droplets. Mass spectrometry revealed that, upon oxidative stress, fatty acids from membrane phospholipids are incorporated into droplet-localized triglycerides. These observations suggest that, upon oxidative stress, vulnerable lipids are rerouted to lipid droplets where they are protected from ROS (Figure 1, right).

There is an interesting twist: although all types of fatty acids relocate to droplets to a similar extent, certain ones are selectively depleted from membranes, namely fatty acids with multiple double bonds, polyunsaturated fatty acids (PUFAs). PUFAs are particularly sensitive to lipid

peroxidation, and thus the pool still in membranes is presumably preferentially damaged and lost. Glial droplets are induced particularly strongly on a high-PUFA diet, as if membranes with more vulnerable lipids require enhanced protection.

But is protection from lipid peroxidation responsible for continued neuroblast proliferation? Although a comprehensive answer will require identifying exactly how neuroblasts are compromised, available evidence supports a mechanistic connection. First, on a high-PUFA diet, glial droplets are particularly important to sustain proliferation. Second, neuroblasts show less peroxidative damage if glia can make lipid droplets. Third, without glial droplets, neuroblasts display increased accumulation of 4-HNE, a product of lipid peroxidation that damages proteins. Exogenous 4-HNE induces ROS, suggesting a feedback loop that amplifies damage and allows its spread across the tissue (Figure 1, right).

In summary, this paper makes a compelling case that lipid droplets play an unanticipated, non-cell-autonomous role in a stem cell niche, allowing stem cells to proliferate even with oxidative stress. Sequestration of harmful lipids and proteins to protect the rest of the cell is a known function of lipid droplets (Walther and Farese, 2012; Welte, 2015); in the proposed model, sequestration protects vulnerable lipids from a noxious environment. Yet how lipid droplets provide a safe haven for PUFAs remains unknown. Is the neutral lipid core simply inaccessible to polar ROS or do co-accumulating lipids or proteins act as antioxidants? Knowledge about composition, cell biology, and biophysics of lipid droplets is exploding; e.g., artificial droplet-like

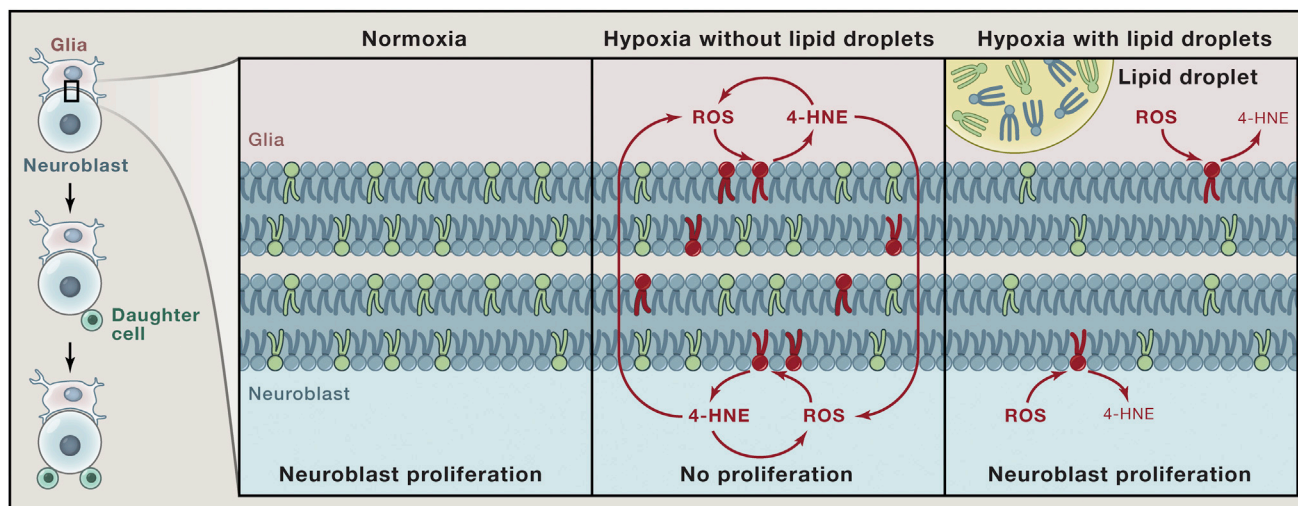


Figure 1. Lipid Droplets Enable Neuroblast Proliferation in Hypoxic Conditions

Left: during normal development, neuroblasts divide asymmetrically, regenerating the neuroblast and generating daughters that will give rise to neurons. Adjacent glial cells provide the stem cell niche for the neuroblasts. Right: a magnified view of the glia/neuroblast interface shows the lipid bilayers that make up the plasma membranes of the glia and neuroblast cells. Lipids containing PUFAs are shown in green; lipids damaged by peroxidation are in red. At normal oxygen levels, neuroblasts keep dividing. Under hypoxic conditions and in the absence of lipid droplets, ROS induces lipid peroxidation in membranes, which in turn gives rise to 4-HNE adducts on proteins. 4-HNE causes more ROS, resulting in an escalating feedback loop. The accumulating damage brings neuroblast proliferation to a halt. When glial lipid droplets are present, hypoxic conditions lead to a redistribution of membrane lipids (including PUFA-containing ones) to lipid droplets where they are protected from ROS-induced damage. Because membranes now contain fewer targets for lipid peroxidation, the feedback loop is dampened, and overall damage remains mild enough for neuroblasts to continue proliferating. It is not fully resolved if hypoxic remodeling lowers the relative concentration of PUFAs in the membrane (e.g., by reducing the ratio of lipids to proteins) or if total membrane surface shrinks.

structures can now be generated in vitro (Thiam et al., 2013). Thus, answers to these questions are within reach.

One central mystery is that all kinds of fatty acids are rerouted from membranes to lipid droplets, not just the most vulnerable PUFAs. Is this bulk movement necessary for protection? It also implies that fewer phospholipids remain in the cellular membranes. Whether this reduces overall membrane surface or alters membrane composition is unknown. Probing membrane properties after hypoxic remodeling will likely provide insightful. It will also be revealing whether membrane lipids from neuroblasts end up in glial droplets or if potential ROS targets are only eliminated from glial membranes.

Oxidative stress is a severe, constant challenge for organisms, contributing to many pathologies as well as to functional decline in aging. I therefore suspect that this newly discovered antioxidant function is deployed widely, especially since lipid droplets are ubiquitous. In mammalian cancer cells, lipid droplets are indeed

known to enhance cell proliferation upon hypoxia (Bensaad et al., 2014). In adult flies and mice, neurodegeneration induces glial lipid droplets in a ROS-dependent manner, but here, presence of droplets correlates with more severe phenotypes (Liu et al., 2015)—in stark contrast to the beneficial effect on larval neuroblasts. This discrepancy might arise from context-specific roles of glial lipid droplets, with different lipid pathways sustaining proliferating neuroblasts and postmitotic neurons.

Since both too little and too much ROS can be harmful, antioxidants levels are generally carefully regulated. This precedent suggests that droplets' antioxidant properties are tunable and that, in patients with reduced or excessive lipid storage, imbalanced ROS may contribute to disease.

REFERENCES

- Bailey, A.P., Koster, G., Guillemier, C., Hirst, E.M.A., MacRae, J.I., Lechene, C.P., Postle, A.D., and Gould, A.P. (2015). *Cell* 163, this issue, 340–353.
- Bensaad, K., Favaro, E., Lewis, C.A., Peck, B., Lord, S., Collins, J.M., Pinnick, K.E., Wigfield, S., Buffa, F.M., Li, J.L., et al. (2014). *Cell Rep.* 9, 349–365.
- Cheng, L.Y., Bailey, A.P., Leivers, S.J., Ragan, T.J., Driscoll, P.C., and Gould, A.P. (2011). *Cell* 146, 435–447.
- Inloes, J.M., Hsu, K.L., Dix, M.M., Viader, A., Masuda, K., Takei, T., Wood, M.R., and Cravatt, B.F. (2014). *Proc. Natl. Acad. Sci. USA* 111, 14924–14929.
- Kis, V., Barti, B., Lippai, M., and Sass, M. (2015). *PLoS ONE* 10, e0131250.
- Liu, L., Zhang, K., Sandoval, H., Yamamoto, S., Jaiswal, M., Sanz, E., Li, Z., Hui, J., Graham, B.H., Quintana, A., and Bellen, H.J. (2015). *Cell* 160, 177–190.
- Thiam, A.R., Farese, R.V., Jr., and Walther, T.C. (2013). *Nat. Rev. Mol. Cell Biol.* 14, 775–786.
- Walther, T.C., and Farese, R.V., Jr. (2012). *Annu. Rev. Biochem.* 81, 687–714.
- Welte, M.A. (2015). *Curr. Biol.* 25, R470–R481.

Winning the Microbial Battle, but Not the War

Hiutung Chu¹ and Sarkis K. Mazmanian^{1,*}

¹Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

*Correspondence: sarkis@caltech.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.050>

An effective immune response leads to rapid elimination of infectious agents, with seemingly little long-term impairment to the host. New findings by Morais da Fonseca et al. reveal that acute infections may result in permanent disruption of tissue homeostasis and immune dysfunction, long after clearance of a pathogen.

Throughout the course of one's lifetime, an individual will experience numerous episodes of acute and chronic microbial infection—including, but not limited to, gastrointestinal and respiratory infections. Exposure to certain pathogens has been shown to have profound, long-lasting consequences on the function of the host immune responses (Medzhitov, 2008, 2010), with some chronic infections proposed as triggers for inflammatory disorders such as multiple sclerosis, lupus, and inflammatory bowel disease (IBD) (Karin et al., 2006; Wucherpfennig, 2001). In contrast, acute infections are typically considered self-limiting without engendering long-term immune consequences to the host, though resolution nevertheless requires an inflammatory response and disruption of tissue homeostasis. In this issue of *Cell*, Morais da Fonseca et al. reveal that a single acute infection with the enteric pathogen *Yersinia pseudotuberculosis* results in a permanent dysregulation of tissue-specific immunity (Morais da Fonseca et al., 2015). Long after clearance of *Yersinia*, the infection appears to cause persistent impairment of the lymphatic network, disrupting molecular and cellular interactions between the gut and mesenteric lymph nodes (MLN), resulting in impaired mucosal immune responses (Figure 1). This exciting new study reveals that acute infections may contribute to long-term disruption of tissue homeostasis, promoting chronic immune dysfunction and disease.

While most interactions between microbes and animals are symbiotic in nature, the immune system has been shaped by an evolutionary “arms race” with pathogens. Examples of chronic infections resulting in persistent inflamma-

tion and localized tissue damage include *Mycobacterium tuberculosis* (tuberculosis), *Helicobacter pylori* (peptic ulcers and gastric cancer), hepatitis B virus and hepatitis C virus (chronic hepatitis and cirrhosis), and *Borrelia burgdorferi* (Lyme disease) (Karin et al., 2006; Wucherpfennig, 2001). However, there are fewer examples illustrating how acute infections may be linked to permanent alterations to host immunity. Rheumatic fever, an inflammatory disease involving the heart and joints, develops following an acute, self-limiting infection of the pharynx with group A *Streptococcus pyogenes*. A pathological consequence of molecular mimicry, the immune response to streptococcal M protein elicits cross-reactivity to human protein antigens, specifically myosin, tropomyosin, and laminin in cardiac myocytes. However, whether acute infections permanently alter general immune function has remained unclear, as has evidence that a pathogen can trigger chronic inflammatory diseases after its clearance by the immune system. Here, Morais da Fonseca et al. report that after an acute gastrointestinal infection with *Y. pseudotuberculosis*, 70% of mice develop chronic mesenteric lymphadenopathy (CL⁺) despite clearance of *Yersinia*. In a series of elegant experiments, the authors reveal that infection-induced tissue damage and remodeling of mesenteric lymphoid tissues promotes impaired integrity of the lymphatic network, as shown by leakage of fluorescent BODIPY from lymphatic vessels into the surrounding mesenteric adipose tissue (MAT). As a result of increased lymphatic permeability, migratory CD103⁺CD11b⁺ DCs were shown to accumulate in the MAT compartment rather than transiting to MLNs as expected (Figure 1), thereby

blunting robust mucosal immune responses. Notably, the defect in lymphatic integrity persisted at least nine months post-*Yersinia* infection, suggesting durable pathologic consequences well after clearance of the pathogen. A similar finding was previously reported where respiratory tract infection with *Mycoplasma pulmonis* led to persistent lymphangiogenesis and reduced transport of antigen presenting cells from the airways to draining lymph nodes (Baluk et al., 2005). Morais da Fonseca and co-workers extend these previous concepts to reveal a mechanism whereby tissue damage caused during acute gastrointestinal infection impairs proper mucosal immune function. Specifically, pathogen-induced reduction in trafficking of migratory gut DCs to the MLN compromises the ability of these cells to promote oral tolerance and protective immunity to oral vaccination.

The authors proceed to show that the microbiota is required to sustain inflammation after resolution of *Yersinia*. Germ-free (GF) mice infected with *Y. pseudotuberculosis* developed lymphadenopathy, similar to specific-pathogen free (SPF) mice. However, migration dynamics of CD103⁺CD11b⁺ DCs from the lamina propria to MLNs remained unaffected in GF mice, suggesting that *Yersinia* infection causes mislocalization of DCs to the MAT via effects on the microbiota. Antibiotic treatment depleting the microbiota in SPF mice was sufficient to reduce cell infiltration into the surrounding MAT, restoring proper DC trafficking and robust mucosal immunity, further implicating the microbiota as a vehicle for sustained immune dysregulation following resolution of acute infection. It is important to note, however, that the

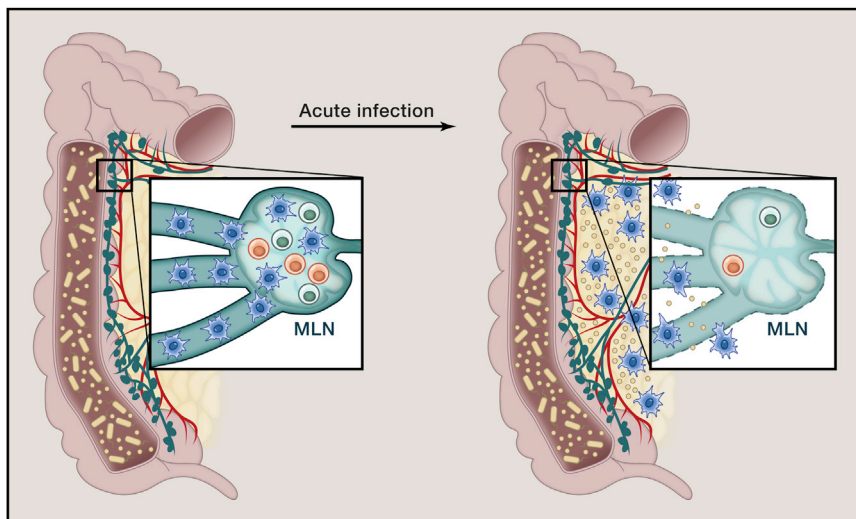


Figure 1. Proposed Model for Durable Inflammatory Responses Following Acute Infection

Under normal conditions (left), migratory CD103⁺ CD11b⁺ DCs recognize and internalize antigens from the gut and traffic to MLNs via lymphatic vessels. In the MLNs, CD103⁺ CD11b⁺ DCs induce TREG and TH17 responses to promote oral tolerance and protective immunity to oral vaccination. Following the clearance of *Yersinia pseudotuberculosis* (right), infection-induced tissue damage leads to increased permeability of the lymphatic network. As a result, migratory DCs accumulate in the MAT compartments rather than trafficking to MLNs, blunting TREG and TH17 responses. The microbiota sustains inflammation following clearance of *Yersinia*, as enhanced exposure to microbial ligands occur due to increased permeability of the lymphatic system. MLN, mesenteric lymph nodes; MAT, mesenteric adipose tissues.

development of chronic lymphadenopathy was not associated with either a bloom of pathogenic bacterial strains (pathobionts) or dysbiosis of the gut microbiota, as lymphadenopathy was not transmissible via co-housing or fecal transplant into uninfected GF mice. Rather, it was likely due to enhanced exposure of the immune system to gut microbial ligands as a result of defective lymphatic integrity, as treatment with bacterial products alone was able to promote and sustain inflammation following enteric infection.

As features of chronic inflammatory diseases typically mirror immune responses to acute infection, identification of a direct link between a causal infectious agent and chronic disease has been challenging. For Crohn's disease, a chronic inflammatory condition of the gastrointestinal tract, many, including Crohn and colleagues, have postulated an inciting infectious

agent may be responsible for the initiation of disease (Crohn et al., 1984; Hugot et al., 2003; Sartor, 2005). Despite extensive research efforts, a causative organism(s) for Crohn's disease has yet to be reproducibly validated, though this issue remains controversial (Greenstein, 2003). Perhaps chronic inflammation observed in Crohn's patients may be the result of persistent effects of a previously encountered infection that caused a long-lasting impairment of the immune system. The study by Morais da Fonseca et al. provides a novel framework to examine how a cleared infection may induce permanent tissue remodeling and altered immune responses, contributing to chronic disease later in life. However, outstanding questions remain despite these comprehensive advances; namely the mechanism by which *Yersinia* promotes chronic lymphadenopathy remains unresolved. What cellular and molecular changes to the

host impair integrity of lymphatic vessels? Are there specific bacterial factors that are required for this outcome? Further, it is unknown why certain populations of mice undergo lymphatic remodeling, whereas others (cage mates) are protected from lymphadenopathy. Not surprisingly, the outcomes uncovered by this new study are unlikely to be specific to *Yersinia* pathogenesis as the authors showed *Toxoplasma gondii* infection, as well as T cell transfer of colitis, are both sufficient to promote chronic lymphadenopathy. These data suggest that chronic inflammatory diseases may be caused by acute infections and are a consequence of a generalized host response. Collectively, the innovative findings from Morais da Fonseca and coworkers reveal an intriguing hypothesis whereby remodeling of the immune system by an acute infection may cause chronic tissue-specific damage and symptoms associated with IBD and autoimmunity and may help explain a possible etiology of various idiopathic inflammatory conditions.

REFERENCES

- Baluk, P., Tammela, T., Ator, E., Lyubynska, N., Achen, M.G., Hicklin, D.J., Jeltsch, M., Petrova, T.V., Pytowski, B., Stackner, S.A., et al. (2005). *J. Clin. Invest.* 115, 247–257.
- Crohn, B.B., Ginzburg, L., and Oppenheimer, G.D. (1984). *JAMA* 251, 73–79.
- Greenstein, R.J. (2003). *Lancet Infect. Dis.* 3, 507–514.
- Hugot, J.P., Alberti, C., Berrebi, D., Bingen, E., and Cézard, J.P. (2003). *Lancet* 362, 2012–2015.
- Karin, M., Lawrence, T., and Nizet, V. (2006). *Cell* 124, 823–835.
- Medzhitov, R. (2008). *Nature* 454, 428–435.
- Medzhitov, R. (2010). *Cell* 140, 771–776.
- Morais da Fonseca, D., Hand, T., Han, S., Gerner, M., Zaretsky, A., Byrd, A., Brenchley, J., Brodsky, I., Germain, R., Randolph, G., et al. (2015). *Cell* 163, this issue, 354–366.
- Sartor, R.B. (2005). *Gut* 54, 896–898.
- Wucherpfennig, K.W. (2001). *J. Clin. Invest.* 108, 1097–1104.

“A Sledgehammer Breaks Glass but Forges Steel”: Bacteria Adhesion Shapes Gut Immunity

Virginia A. Pedicord¹ and Daniel Mucida^{1,*}

¹Laboratory of Mucosal Immunology, The Rockefeller University, 1230 York Ave, New York, NY 10065, USA

*Correspondence: mucida@rockefeller.edu
<http://dx.doi.org/10.1016/j.cell.2015.09.040>

Gut bacteria are known to affect immune cell development, but most intestinal lymphocytes have no direct contact with luminal bacteria. Two studies by Atarashi et al. and Sano et al. shed light on how bacterial adhesion can cue intestinal epithelial cells to direct differentiation of gut T cells.

The mammalian intestine houses millions of bacteria, which under homeostatic conditions cooperate with host cells to facilitate normal metabolism, immune system development, and pathogen resistance. This is largely conducted in a contact-independent manner via bacterial metabolites, since the majority of intestinal bacteria are non-adherent and do not cross the dense mucous layer that overlies the epithelium. For instance, short-chain fatty acids produced by certain bacteria in the lumen of the colon readily diffuse across epithelial cell membranes to be directly utilized by epithelial cells themselves or to instruct differentiation of gut immune cells, particularly regulatory T cells (Smith et al., 2013). In addition, unique microbial molecular patterns, such as fragments of bacterial cell walls, are sensed by immune cells to trigger secretion of soluble factors and amplification of immune responses at sites of infection. At the interface between the microbe-rich lumen and the underlying immune-rich lamina propria, intestinal epithelial cells may therefore serve as major transmitters of microbe-derived immunomodulatory signals in the gut. In this issue of *Cell*, Atarashi et al. (2015) and Sano et al. (2015) now elucidate how bacterial adhesion to intestinal epithelial cells can direct the functional differentiation of gut Th17 cells.

Production of the cytokine IL-17, the defining characteristic of Th17 cells, has been shown to occur in CD4⁺ T cells in protective responses against several extracellular mucosal pathogens such as *Candida albicans* and *Citrobacter rodentium* (Conti et al., 2009; Mangan et al., 2006). Indigenous gut bacteria are required to induce steady-state Th17 dif-

ferentiation as evidenced by the near absence of Th17 cells in germ-free animals (Ivanov et al., 2009). Through systematic colonization of germ-free mice with various bacterial isolates, Gaboriau-Routhiau et al. found commensal segmented filamentous bacteria (SFB) to be sufficient to activate various effector CD4⁺ T cell responses in the gut, particularly Th17 responses (Gaboriau-Routhiau et al., 2009). By comparing the microbiota of Th17-sufficient mice (Taconic) to that of naturally Th17-insufficient mice (Jackson), Ivanov and coworkers simultaneously identified SFB as the major microbial determinant of Th17 induction in conventional Taconic C57BL/6 mice (Ivanov et al., 2009). However, it has remained unclear what characteristic of these microbes elicits IL-17 expression and how Th17 cells arise in the absence of pathogenic infection.

In this new study, Sano et al. (2015) utilized SFB-specific transgenic T cells and MHCII tetramers to track the maturation of SFB-specific CD4⁺ T cells temporally along the gastrointestinal tract after acute colonization. They observed that, while “poised” transgenic T cells expressing the Th17 master regulator ROR γ t could be found throughout the gut, they were most abundant in the small intestine ileum. The ileum was also the location with the highest level of SFB colonization, and it was only here that transgenic T cells robustly expressed IL-17. In their seminal study, Gaboriau-Routhiau et al. previously noted that the most robust cytokine induction occurred in colonization conditions that included SFB adherence to Peyer’s patches (Gaboriau-Routhiau et al., 2009). Now, in elegant complementary studies, Atarashi et al. (2015) show

this preferential IL-17 induction to be highly dependent on enhanced SFB adhesion to the small intestine epithelium, particularly the follicle-associated epithelium. By inoculating mice with a rat strain of SFB that is equally efficient at colonizing but unable to adhere to small intestine epithelial cells, they demonstrated that SFB colonization itself was not sufficient to induce Th17 cells in the absence of adhesion. They went on to establish that other adherent microbes (*C. rodentium*, EHEC, and *C. albicans*) are also able to promote IL-17 production and that this property is lost when these microbes are rendered non-adherent.

Ivanov et al. previously showed that Th17 induction was associated with increased expression of all three isoforms of serum amyloid A (SAA) in the terminal ileum (Ivanov et al., 2009). However, the vital role of epithelial cell-derived SAA expression in Th17 differentiation has only now been revealed. Using SAA-deficient animals and an in vitro Th17 differentiation system, Sano et al. (2015) determined that SAA was both necessary and sufficient to directly promote IL-17 production and proliferation in ROR γ t⁺ CD4⁺ T cells. Atarashi et al. (2015), on the other hand, described SAA-dependent production of Th17-promoting cytokines by CD11c⁺ myeloid cells and an additional requirement for adhesion-elicited epithelial cell production of reactive oxygen species. Although neither study focused on the identification of the myeloid lineage responsible for these effects, other recent studies have uncovered a critical role for monocyte-derived CX₃CR1⁺ cells in the generation of SFB-specific Th17 responses (Panea et al., 2015). Additionally, the Littman group

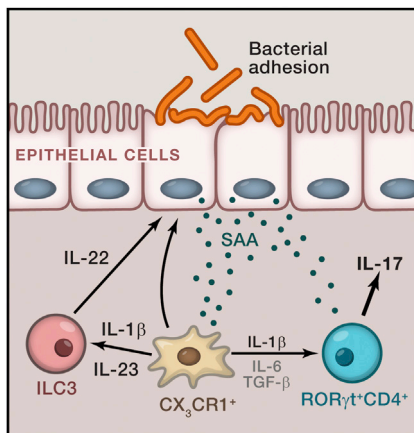


Figure 1. Bacterial Adhesion to Intestinal Epithelial Cells Forges Gut T Cell Function

In response to attachment of SFB or other adherent bacteria (the “hammer”), intestinal epithelial cells upregulate and release serum amyloid A (SAA). SAA boosts myeloid cell (presumably CX₃CR1⁺ mononuclear antigen-presenting cells) production of IL-1b and perhaps IL-23. IL-1b can amplify SAA expression by signaling directly to the IECs, and IL-1b and IL-23 promote production of IL-22 in type 3 innate lymphoid cells (ILC3). IL-22, in turn, reinforces SAA expression by IECs. This circuitry promotes IL-17 expression in “poised” RORγt⁺ CD4⁺ T cells.

demonstrated that CX₃CR1⁺ cell-derived IL1-β and IL-23 induces ILC3 production of IL-22 (Longman et al., 2014). While both Sano et al. (2015) and Atarashi et al. (2015) identify a requirement for SFB-induced ILC3 production of IL-22 in potentiating SAA expression by epithelial cells, it is still unclear whether IL-22 is sufficient to prompt SAA-mediated IL-17 expression. It is therefore plausible that CX₃CR1⁺ cells are uniquely reactive to SAAs, possibly by preferential expression

of SAA receptors or lipoprotein particle receptors, and are thereby specialized to amplify epithelial cell SAA expression (via IL-1β) and contribute to Th17 induction. These Th17-inducing effects of SAAs may also be balanced by a possible regulatory role of retinol-binding SAAs (Derebe et al., 2014), opening up the possibility that the nature of co-factors carried by SAA may influence regulatory versus effector T cell function in the gut.

Much like the sledgehammer forging steel (Leon Trotsky), it is possible that epithelial cells are exquisitely responsive to cytoplasmic perturbations induced by physical and mechanical cues. Hence, bacterial adhesion (the “hammer,” Figure 1) may shape downstream signaling, such as Stat3 phosphorylation (Sano et al., 2015) or C/EBPδ upregulation (Atarashi et al., 2015), via mechanotransduction mechanisms. Although neither the intracellular signaling pathways triggered upon bacterial adhesion nor the exact upstream events that lead to SAA secretion are fully understood, these studies define a new pathway of epithelial cell sensing of intestinal microbes and modulation of adaptive immune responses, highlighting the importance of local environmental cues in licensing effector cytokine production. This could have important implications to the physiological functions of Th17 cells in local control of pathogenic infections and systemic autoimmune responses.

ACKNOWLEDGMENTS

We thank Thiago L. Carvalho and Lillie Cohn for their helpful suggestions with the manuscript.

REFERENCES

- Atarashi, K., Tanoue, T., Ando, M., Kamada, N., Nagano, K., Narushima, S., Suda, W., Imaoka, A., Setoyama, H., Nagamori, T., et al. (2015). *Cell* 163, this issue, 367–380.
- Conti, H.R., Shen, F., Nayyar, N., Stocum, E., Sun, J.N., Lindemann, M.J., Ho, A.W., Hai, J.H., Yu, J.J., Jung, J.W., et al. (2009). *J. Exp. Med.* 206, 299–311.
- Derebe, M.G., Zlatkov, C.M., Gattu, S., Ruhn, K.A., Vaishnav, S., Diehl, G.E., MacMillan, J.B., Williams, N.S., and Hooper, L.V. (2014). *eLife* 3, e03206.
- Gaboriau-Routhiau, V., Rakotobe, S., Lécuyer, E., Mulder, I., Lan, A., Bridonneau, C., Rochet, V., Pisi, A., De Paepe, M., Brandi, G., et al. (2009). *Immunity* 31, 677–689.
- Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., et al. (2009). *Cell* 139, 485–498.
- Longman, R.S., Diehl, G.E., Victorio, D.A., Huh, J.R., Galan, C., Miraldi, E.R., Swaminath, A., Bonneau, R., Scherl, E.J., and Littman, D.R. (2014). *J. Exp. Med.* 211, 1571–1583.
- Mangan, P.R., Harrington, L.E., O’Quinn, D.B., Helms, W.S., Bullard, D.C., Elson, C.O., Hatton, R.D., Wahl, S.M., Schoeb, T.R., and Weaver, C.T. (2006). *Nature* 441, 231–234.
- Panea, C., Farkas, A.M., Goto, Y., Abdollahi-Roodsaz, S., Lee, C., Koscsó, B., Gowda, K., Hohli, T.M., Bogunovic, M., and Ivanov, I.I. (2015). *Cell Rep.* 12, 1314–1324.
- Sano, T., Huang, W., Hall, J.A., Yang, Y., Chen, A., Gavzy, S.J., Lee, J.-Y., Ziel, J., Miraldi, E.R., Domingos, A.I., et al. (2015). *Cell* 163, this issue, 381–393.
- Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly-Y, M., Glickman, J.N., and Garrett, W.S. (2013). *Science* 341, 569–573.

Fleeting Amyloid-like Forms of Rim4 Ensure Meiotic Fidelity

Alice Flynn Ford^{1,2} and James Shorter^{1,2,*}

¹Department of Biochemistry and Biophysics

²Neuroscience Graduate Group

Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA 19104, USA

*Correspondence: jshorter@mail.med.upenn.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.049>

Berchowitz et al. establish that transient amyloid-like forms of Rim4, a yeast RNA-binding protein with a predicted prion domain, translationally repress cyclin *CLB3* in meiosis I, thereby ensuring homologous chromosome segregation. These findings suggest that prion domains might enable formation of tightly regulated amyloid-like effectors in diverse functional settings.

Meiosis yields haploid gametes from diploid progenitors via two consecutive rounds of chromosome segregation, meiosis I and II, following a single S phase. During meiosis I, homologous chromosomes segregate, whereas in meiosis II sister chromatids separate. Proper meiotic progression requires translational regulation of specific cyclins, such as the B-type cyclin, *CLB3* in budding yeast, which is repressed during meiosis I but translated during meiosis II (Carlile and Amon, 2008). Aberrant expression of *CLB3* during meiosis I prevents the essential segregation of homologous chromosomes and triggers premature sister chromatid separation (Carlile and Amon, 2008). In this issue of *Cell*, an elegant study by Berchowitz et al. (2015) now establishes that Rim4, a meiosis-specific RNA-binding protein (RBP) with a predicted prion domain (Figure 1A) (Alberti et al., 2009), forms functional amyloid-like conformers that directly repress *CLB3* translation during meiosis I (Berchowitz et al., 2015).

Rim4 is a 713-residue RBP with three N-terminal RNA-recognition motifs (RRMs) and a C-terminal low-complexity (LC) domain predicted to be predominantly disordered, which harbors a putative prion domain (Figure 1A) (Alberti et al., 2009). Prion domains have an unusual, low-complexity amino acid composition comprised of predominantly uncharged polar amino acids (particularly glutamine, asparagine, and tyrosine) and glycine (Alberti et al., 2009; King et al., 2012). In canonical yeast prion proteins, the prion domain can switch from an

intrinsically disordered conformation to an infectious cross- β conformation (Alberti et al., 2009; King et al., 2012). In the present study, the authors showed that in isolation, Rim4 spontaneously assembled into β sheet-rich fibrils, and that Rim4 fibrillization was RNA-independent and required the C-terminal LC domain (Berchowitz et al., 2015). Moreover, in vivo, the predicted prion domain plus C-terminal determinants in the LC domain of Rim4 mediated formation of stable, SDS-resistant aggregates during premeiotic S phase, which sequestered specific mRNAs and repressed cyclin *CLB3* translation exclusively during meiosis I (Figure 1B) (Berchowitz et al., 2015). Indeed, Rim4 variants bearing C-terminal deletions in the LC domain were unable to form amyloid-like species or elicit translational repression in this context (Berchowitz et al., 2015). Remarkably, at the onset of meiosis II, increased Ime2 kinase activity triggers rapid degradation of amyloid-like Rim4 to enable *CLB3* translation and sister chromatid segregation (Figure 1B) (Berchowitz et al., 2015). Collectively, these data suggest that the formation and elimination of amyloid-like Rim4 is exquisitely timed and tightly orchestrated to enable a specific translational repression modality critical for meiosis I.

Is this mode of regulation specific to yeast? The answer is probably no. In fact, humans possess four meiosis-specific RBPs with predicted prion domains, DAZ1-4 (King et al., 2012). Moreover, a related RBP, DAZL (DAZ-like), is crucial for gametogenesis in mice. In the current

work, the authors observed that DAZL forms SDS-resistant aggregates in cells undergoing meiosis but not in non-meiotic tissue (Berchowitz et al., 2015). Thus, similar to Rim4 in yeast, an amyloid-like form of DAZL, and perhaps also DAZ1-4, might play a role in translational regulation during mammalian gametogenesis.

Rim4 is not the first example of amyloid-like species regulating mRNA translation. However, its repression of *CLB3* translation stands in stark contrast to the translational activation of specific transcripts sparked by amyloid-like forms of a neuronal cytoplasmic polyadenylation element binding protein from *Aplysia californica* (ApCPEB) (Shorter and Lindquist, 2005; Si et al., 2010). ApCPEB also harbors a prion domain, which drives assembly of amyloid-like forms of ApCPEB in response to specific neurotransmitter cues that stimulate a translational program that enables synaptic robustness and long-term facilitation (Shorter and Lindquist, 2005; Si et al., 2010). Like Rim4, the amyloid-like form of ApCPEB engages specific RNAs more avidly (Berchowitz et al., 2015; Shorter and Lindquist, 2005). This shared ability of amyloid-like forms of Rim4 and ApCPEB to modulate mRNA translation suggests that this type of regulation may be more widespread than presently appreciated. Remarkably, ~1% of human proteins harbor a predicted prion domain, including at least ~50 RBPs (Kim et al., 2013; King et al., 2012; Li et al., 2013). The pervasiveness of these domains is cause to ask—is

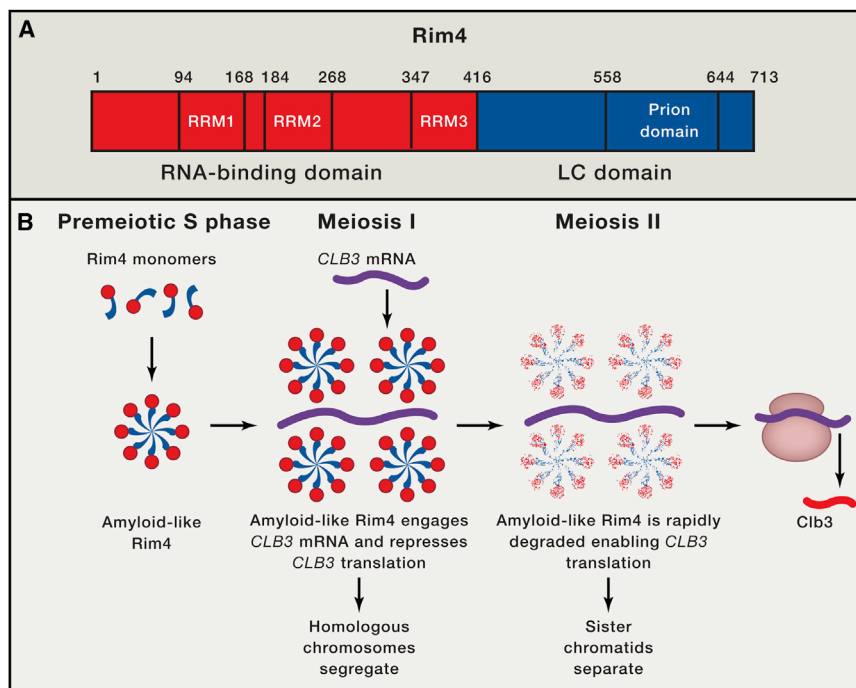


Figure 1. Amyloid-like Forms of Rim4 Repress *CLB3* Translation During Meiosis I

(A) Domain architecture of Rim4. Rim4 is divided into an N-terminal RNA-binding domain (red) harboring three RNA-recognition motifs (RRM1-3) and a C-terminal low-complexity (LC) domain (blue) harboring a predicted prion domain.

(B) Rim4 expression is induced in premeiotic S phase, and the LC domain of Rim4 (blue) enables assembly into amyloid-like species. Upon entry into meiosis I, amyloid-like Rim4 represses the translation of cyclin *CLB3* by binding the 5'UTR of the *CLB3* mRNA (purple). Repression of *CLB3* translation enables segregation of homologous chromosomes. Upon transition to meiosis II, Ime2 kinase activity increases and elicits rapid degradation of amyloid-like Rim4, which relieves inhibition of *CLB3* translation thereby facilitating sister chromatid separation.

the formation of amyloid or amyloid-like states a widely employed mechanism of cellular regulation underlying translational control and perhaps a spectrum of other molecular processes?

It is important to note that amyloid or amyloid-like conformers are not the only structures encoded by predicted prion domains. Indeed, prior to populating self-templating amyloid-like states, these intrinsically disordered regions can undergo concentration-dependent liquid demixing phase transitions. These transitions enable partitioning of various dynamic compartments with liquid- or gel-like properties, including P bodies, stress granules, and paraspeckles (Hennig et al., 2015; Li et al., 2013; Patel et al., 2015). Despite the clear utility of both preamyloid and amyloid-like conformations encoded by predicted prion domains, this type of

domain is a double-edged sword. Indeed, RBPs with predicted prion domains, including FUS, TDP-43, hnRNPA1, and hnRNPA2, have emerged as causative agents in several fatal neurodegenerative disorders such as amyotrophic lateral sclerosis and frontotemporal dementia (Kim et al., 2013; King et al., 2012; Li et al., 2013). Here, the predicted prion domains enable RBP misfolding into deleterious oligomeric and fibrillar structures that likely underpin disease, and disease-linked mutations can accelerate these misfolding events (Kim et al., 2013; King et al., 2012; Li et al., 2013; Patel et al., 2015). In particular, an aberrant phase transition from liquid- or gel-like states to more intractable solid phases mediated by the prion domain is likely a key pathogenic event (Li et al., 2013; Patel et al., 2015).

A critical goal is to accurately understand what primary sequence elements and auxiliary factors determine whether predicted prion domains assemble into functional, beneficial structures or deleterious, misfolded conformers that are pathogenic. Such understanding will empower the design of therapeutics to prevent or reverse detrimental RBP misfolding trajectories. In this context, a mechanistic understanding of the developmentally regulated formation and elimination of amyloid-like Rim4 is likely to be extremely valuable. For example, the mechanism by which amyloid-like Rim4 is rapidly degraded during meiosis II remains uncharted. It is not clear how increased Ime2 kinase activity elicits this degradation process or whether it requires protein disaggregases (e.g., Hsp110 or Hsp104), proteasomes, or autophagy. Delineation of this temporally restricted clearance pathway could inform effective strategies to eliminate pathogenic amyloid conformers.

REFERENCES

- Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). *Cell* 137, 146–158.
- Berchowitz, L.E., Walker, M.R., Kabachinski, G., Carlile, T.M., Gilbert, W.V., Schwartz, T.U., and Amon, A. (2015). *Cell* 163, this issue, 406–418.
- Carlile, T.M., and Amon, A. (2008). *Cell* 133, 280–291.
- Hennig, S., Kong, G., Mannen, T., Sadowska, A., Kobelke, S., Blythe, A., Knott, G.J., Iyer, K.S., Ho, D., Newcombe, E.A., et al. (2015). *J. Cell Biol.* 210, 529–539.
- Kim, H.J., Kim, N.C., Wang, Y.D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., et al. (2013). *Nature* 495, 467–473.
- King, O.D., Gitler, A.D., and Shorter, J. (2012). *Brain Res.* 1462, 61–80.
- Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). *J. Cell Biol.* 201, 361–372.
- Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoykov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). *Cell* 162, 1066–1077.
- Shorter, J., and Lindquist, S. (2005). *Nat. Rev. Genet.* 6, 435–450.
- Si, K., Choi, Y.B., White-Grindley, E., Majumdar, A., and Kandel, E.R. (2010). *Cell* 140, 421–435.

A Biological Imitation Game

Christof Koch^{1,*} and Michael A. Buice¹

¹Allen Institute for Brain Science, Seattle, WA 98109, USA

*Correspondence: christofk@alleninstitute.org

<http://dx.doi.org/10.1016/j.cell.2015.09.045>

The digital reconstruction of a slice of rat somatosensory cortex from the *Blue Brain Project* provides the most complete simulation of a piece of excitable brain matter to date. To place these efforts in context and highlight their strengths and limitations, we introduce a Biological Imitation Game, based on Alan Turing's Imitation Game, that operationalizes the difference between real and simulated brains.

The much awaited *opus magnum* of the “Blue Brain Project” (“BBP”) is appearing in this issue of *Cell* (Markram et al., 2015). The brain child, so to speak, of the electrophysiologist Henry Markram, the BBP is a large-scale, international collaborative effort that seeks to simulate a synthetic brain using biologically realistic models of synapses and nerve cells. The BBP is an exceptional (neuro)-engineering effort based on a bold and long-term vision. To be successful, it requires meticulous attention to detail, a highly interdisciplinary team, and large-scale and stable resources. Financially, it is backed by the Swiss government, with more recent contributions by the pan-European Human Brain Project.

Initiated a decade ago, the ambitious aims of the BBP are to amass all relevant knowledge pertaining to the mammalian brain; to distill the associated information into an integrated, standardized, and open-access database; and to endow this massive yet static entity with life by simulating spontaneous and sensory-driven synaptic activity and the associated cellular electrical responses using partial differential equations. The *Cell* paper describes a first-draft digital reconstruction of a section of rat somatosensory cortex. It is, without doubt, the most complete simulation to date of a piece of excitable brain matter.

The utility of such massive simulations has been much debated—in particular, as they relate to the ill-defined goal of “understanding the brain” that is on the masthead of the Human Brain Project and the other large-scale brain projects initiated over the last 3 years (Kandel et al., 2013). So before we dive into the details, let us step back and consider how best to think about this category-defying model.

Introducing a Biological Imitation Game

In an effort to circumvent the question “can machines think”?, the mathematician Alan Turing (Turing, 1950) introduced “The Imitation Game” in which an interrogator is tasked with determining the identity of two examinees as “male” or “female.” This game is imagined to be played by human beings. Turing introduced as a measure of the thinking ability of machines the length of time the Interrogator needed to play before realizing that one of the other players had been replaced by a machine. The longer this takes, the more this machine could be argued to operationally behave like a thinking human.

Let us imagine our own Imitation Game, in which a lead investigator running a lab is recording from neocortical neurons in a

rat but has also devised a sophisticated computer model of its electrical behavior. Being quite busy with grant writing and other administrative tasks, she rarely enters the lab, relying instead on reports from her students and postdocs.

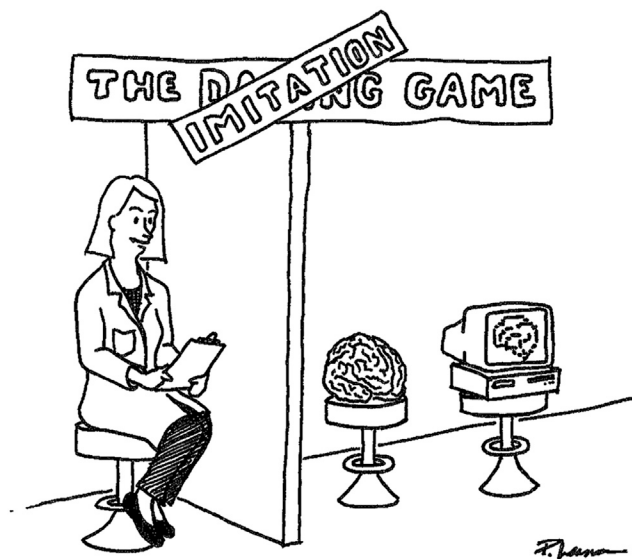
After some discussion, they agree on a protocol X—say, stimulate sensory input into somatosensory cortex while recording from layer 5 pyramidal neurons. The experimentalists in her lab judiciously carry out these manipulations and bring her a set of plots—extra-cellular recordings, current source density plots, and so on. Their computational colleagues do likewise and generate the same type of graphs. Let's call these, respectively, A(X) and B(X).

For the purposes of this operational definition, we assume that the investigator does not know whether A or B is the experimental data, taken here as ground truth. Both look superficially similar. To distinguish between the real data, A(X), and the simulated data, B(X), the investigator must devise clever experiments for her group to carry out on both, probing for the simulation's failure modes. Clearly, the longer that it will take the investigator, given her accumulated knowledge and insights about real brains, to confidently identify the provenance of the data, the more the model can be said to mimic the real system. We take as a measure of the quality of the simulation the length of time it takes her to confidently identify the provenance of the data and the level of sophistication of question required to do so.

Thus, the question of how good any one computer simulation of reality is can be replaced by an operational measure of how long an expert can be fooled by the simulation. As the model becomes more and more sophisticated, this will take longer and longer. In the limit of a completely faithful digital simulacrum, reliably judging which is real and which is synthetic will become impossible. Note that, in the spirit of Turing, our operational definition bypasses eristic discussions of what is meant by “understanding” the brain.

Looking Under the Hood

We will now summarize the approach taken by the BBP team. The model simulates a segment of 2-week-old rat somatosensory cortex, 2 mm tall and 210 μ m in radius, about the size of a cortical column. The individual nerve cells are derived from the Markram laboratory's heroic efforts over almost two decades to record and label more than 14,000 neurons in slices taken from



Bachelor #1: What is your favorite neurotransmitter to release during spreading waves? Illustration by Phil Lesnar, courtesy of the Allen Institute.

young male Wistar rats (e.g., [Markram et al., 1997](#) and [2004](#)). The dendritic trees and local axonal morphologies of 1,009 of these cells are completely digitally reconstructed in 3D. The cells fall into 55 morphological classes, as determined by the layer in which the cell bodies were located and their morphology: 42 GABAergic and 13 glutamergic types. This is roughly in line with the literature ([Harris and Shepherd 2015](#)). The electrical properties of these neurons are inferred by recording from the cell bodies of 3,900 neurons in young rat slices. Using a standard electrophysiological protocol, 11 types of firing patterns—ten associated with inhibitory and one (continuous adapting) with excitatory cells—can be distinguished. From a combined dataset of 511 morphologically and electro-physiologically characterized neurons, the BBP team infers 207 morpho-electrical types (207 out of a possible $55 \times 11 = 605$ cell types).

Next, the team populates their column with 31,320 neurons drawn from these 207 types, with multiple clones of each cell type randomly and independently spatially distributed across five layers (L1, L2/3, L4, L5, and L6) in accordance with measured cell densities and immunohistochemical staining.

The electrical behavior of each neuron is modeled using a classical Hodgkin-Huxley formalism extended to include 13 voltage- and calcium-dependent conductances inserted into the membrane at the cell body and the proximal and distal dendritic tree ([Hay et al., 2011](#)).

At this stage, these neurons with their thick dendrites and thin non-myelinated axons are scattered throughout the cylindrical volumes, a dense jungle with roots, flowering plants, bushes, trees, and their canopy intertwined. Both cortex and forest share a predominantly vertical organization. They differ as only the former have highly specialized junctions among their members, chemical synapses that convert electrical activity in presynaptic axonal boutons into neurotransmitter release

and back into electrical activity at postsynaptic sites on spines, dendrites, or cell bodies.

The Connectomics Algorithm

But where to place the all-important synapses? Here, we come to the beating heart of the BBP, an algorithmic approach to reconstruct synaptic connections (described in detail in a companion paper [[Reimann et al., 2015](#)]). Whenever an axonal profile comes into close encounter with a dendrite, a potential synapse is marked. Such a potential synapse is the child of chance, based on the probabilities that an axonal bouton finds itself near a dendrite. This is known as Peter's rule, stipulating that axons seek out dendrites randomly—the higher the bouton density and the larger the dendrite, the bigger the connection probability. By this measure, most neurons connect to most other neurons (there are, after all, close to one billion synapses in each mm^3 of gray matter). Peters' rule places a hard constraint on the connectivity, in the sense that both axons and dendrites need to be simultaneously present for a synapse to form.

[Reimann and colleagues \(Reimann et al., 2015\)](#) know from light- and electron-microscopic data derived by a great many labs, including their own, that any one functional connection among two neurons is implemented by many anatomical synapses. Put differently, if two cortical neurons are connected, they form anywhere between a handful and perhaps 20 anatomical synapses; only rarely does a single synapse connect two neurons. The observation that simple axonal-dendritic proximity—which would predict a majority of singleton connections—does not explain connectivity highlights the limit of Peters' rule. It also points to the existence of learning rules that prune and grow synaptic connections.

The BBP captures such observed anatomical and morphological constraints in a probabilistic connectomics algorithm that prunes the vast number of spatial appositions, leaving a much smaller number of actual anatomical synapses: 638 million potential synapses yield 37 million actual synapses that form 8.1 million connections, with 3.6 excitatory and 13.9 inhibitory synapses per connection (all values given here are averages; the paper gives statistical distributions). These data are compatible with a recent electron-microscopic reconstruction of a sliver of cortical tissue ([Kasthuri et al., 2015](#)).

The physiology of these 37 million synapses is based on extrapolated data from paired-cell recordings, including their all-important short-term facilitating or depressing dynamic behavior. 207 types of neurons could, in principle, be connected using $207 \times 207 = 42,849$ different types of synapses. Unfortunately, for the vast majority of synaptic types, no experimental recordings are available, and reasonable inter- and extrapolations and assumptions must come to the rescue.

Booting up the Virtual Slice

With synaptic input in place, fortified by somatic depolarization of cells to facilitate their firing (mimicking arousing neuromodulator substances active under *in vivo* conditions) and some spontaneous synaptic release, the transmembrane potential of each cell evolves according to the nonlinear cable equation. This is a partial differential equation that describes how the membrane

potential evolves in thin and elongated neuronal processes whose membrane properties are described by capacitive and conductive processes (Koch 1999). The numerical simulation uses a parallel version of the NEURON software that has been instrumental to the field (Carnevale and Hines 2006), running on an IBM Blue Gene/Q supercomputer at the Swiss National Supercomputing Center (CSCS) in Lugano.

Powering up this digital slice yields reasonable-looking firing behavior and slow oscillatory bursts at 1 Hz. The simulation goes through some finger exercises, demonstrating its neurobiological authenticity. The overall network activity is very sensitive to the extracellular calcium concentration $[Ca^{2+}]_o$ and the level of somatic depolarization of all neurons. $[Ca^{2+}]_o$, varying between 1 mM under in vivo and 2 mM under in vitro conditions affects the probability of synaptic release. In effect, $[Ca^{2+}]_o$ controls the network's susceptibility to synaptic perturbation, while the depolarization level controls the spontaneous firing of the neurons. Manipulating these variables flips the network between two qualitatively different dynamic regimes. Under in vitro conditions ($[Ca^{2+}]_o = 2$ mM and less somatic depolarization), the network is highly responsive to perturbations, producing stereotypical and synchronized responses, such as spreading waves (Figure 14 in Markram et al., 2015), while under in vivo conditions ($[Ca^{2+}]_o$ closer to 1 mM and greater somatic depolarization), responses are graded and asynchronous. This transition is quite sharp and is mediated by differential changes in the excitatory to inhibitory synaptic balance.

Given the very large number of approximations and extrapolations that underlie the BBP model, the fact that neurons do not erupt into either a paroxysm of epileptic discharge nor descend into a coma-like state of electrical silence but act, to a first order, as neurons in slices do, is a remarkable achievement. This is an important initial success in our Imitation Game, for the very reason that it allows the game to continue. Our investigator can begin asking tougher questions (of course, it also begs the question of whether all of these details were necessary in the first place, to which we will return later on).

Pushing further, the BBP team highlights aspects of in vivo activity that are recapitulated by the model—in particular, the extent to which groups of neurons are active in a correlated manner (for instance, when spikes fire near simultaneously), have no particular temporal relationship, or are anti-correlated. Indeed, the digital simulacrum exhibits the same correlational structure as predicted on theoretical grounds and observed in rodent cortex by Renart et al. (2010). In particular, activities in excitatory and inhibitory neurons track each other, generating negative correlations in synaptic structure that cancel strong shared input. Similarly, the simulation shows repeating triplet structures in the synchronous regime that do not appear in the asynchronous regime of low calcium. Finally, a recent experimental study of Okun et al. (2015) establishes the existence of many highly correlated “chorister” neurons and a few isolated “soloists.” The former are neurons whose activity is tightly correlated with the average activity of the network that they are embedded in, while the latter appear to actively fire independently of the rest of the network. The simulation recapitulates this observation as well.

These demonstrations are comforting validations of the model in terms of our Imitation Game. Of course, the transition observed by the BBP team is clearly describable by simple models (there is a long history, with notable contributions from Wilson and Cowan (1972) and Amari (1975) of applied mathematics devoted to the solution space of so-called “neural field theories,” demonstrating complex behavior as neural characteristics are varied (see Bressloff, 2014), including spreading waves (Beurle, 1956).

The most interesting observation of the present study is the sensitivity of network behavior to extracellular calcium concentration and that propagating waves may be an artifact of too high $[Ca^{2+}]_o$ levels (which would not, however, explain the waves seen in Xu et al., 2007). Yet the enormous effort of the BBP to carefully reconstruct details of the cortical circuit are probably irrelevant to this prediction, as the dependency on extracellular calcium appears to enter through an empirical model of synaptic neurotransmitter release and not through any network dynamics. In the same vein, the mechanism of Renart et al. (2010) was obtained in a model of conductance-based integrate-and-fire neurons. The additional detail added by the BBP does not affect the mechanism.

Uncovering the Limits of the Model

In the Imitation Game, there is one cunning strategy that is guaranteed to lead to the denouement of the impostor. Simply probe further and further into the micro-structure: while it is not clear whether there is any “ultimate” level of reality—elementary particles such as the Higgs boson, fields, or superstrings (depending on the energy invested in probing the micro-structure)—simulations abruptly bottom out. To wit, the investigator could ask her lab members to resolve fine details of the macroscopic Na^+ , K^+ , Ca^{2+} , and Cl^- membrane currents. If appropriate electrodes and amplifiers are used, the recordings will uncover discrete, microscopic, and stochastic ionic channels, while the current BBP model stops with the continuous and deterministic Hodgkin-Huxley currents. Bingo!

And therein lies an important lesson. If the real and the synthetic can't be distinguished at the level of firing rate activity (even though it is uncontroversial that spiking is caused by the concerted action of tens of thousands of ionic channel proteins), the molecular level of granularity would appear to be irrelevant to explain electrical activity. Teasing out which mechanisms contribute to any specific phenomena is essential to what is meant by understanding.

Markram et al. claim that their results point to the minimal datasets required to model cortex. However, we are not aware of any rigorous argument in the present triptych of manuscripts (Markram et al., 2015; Ramaswamy et al., 2015; Reimann et al., 2015), specifying the relevant level of granularity. For instance, are active dendrites, such as those of the tall, layer 5 pyramidal cells (Hay et al., 2011), essential? Could they be removed without any noticeable effect? Why not replace the continuous, macroscopic, and deterministic HH equations with stochastic Markov models of thousands of tiny channel conductances? Indeed, why not consider quantum mechanical levels of descriptions? Presumably, the latter two avenues have not been chosen because of their computational burden and the

intuition that they are unlikely to be relevant (Strassberg and DeFelice, 1993; Koch and Hepp, 2006). The Imitation Game offers a principled way of addressing these important questions: only add a mechanism if its impact on a specific set of measurables can be assessed by a trained observer.

Consider the problem of numerical weather prediction and climate modeling, tasks whose physico-chemical and computational complexity is comparable to whole-brain modeling. Planet-wide simulations that cover timescales from hours to decades require a deep understanding of how physical systems interact across multiple scales and careful choices about the scale at which different phenomena are modeled. This has led to an impressive increase in predictive power since 1950, when the first such computer calculations were performed (Bauer et al., 2015). Of course, a key difference between weather prediction and whole-brain simulation is that the former has a very specific and quantifiable scientific question (to wit: “is it going to rain tomorrow?”). The BBP has created an impressive initial scaffold that will facilitate asking these kinds of questions for brains.

What is also noteworthy is that, with the publication of this paper, the BBP has publicly released a treasure trove of morphological, electrical, and connectional data collected over the last years by members of the Markram lab that was used in this model. For more details, see the companion manuscript (Ramswamy et al., 2015). The experience of the Allen Institute (www.brain-map.org) has demonstrated the impact of releasing all relevant data and models soon after they pass an internal quality control stage (usually several years prior to publication of any associated results). The release of the massive database of the BBP is thus a very positive development.

It is clear that physically detailed, whole-brain simulations enable neuroscientists to answer specific questions that are difficult to address experimentally. The BBP has provided a powerful tool in this regard. Such simulations may be essential to develop therapeutics for brain-based diseases. Whole-brain models can also be said to encapsulate the present state of knowledge, as per the Imitation Game. However, any simulation-gained knowledge must be supported and complemented by theories that, by isolating the relevant variables, enable us to deeply understand the most organized and highly excitable piece of matter in the known universe.

REFERENCES

- Amari, S. (1975). *Biol. Cybern.* 17, 211–220.
- Bauer, P., Thorpe, A., and Brunet, G. (2015). *Nature* 525, 47–55.
- Beurle, R.L. (1956). *Phil. Trans. Roy. Soc. B.* 240, 55–94.
- Bressloff, P.C. (2014). *Waves in Neural Media: From Single Neurons to Neural Fields* (Springer).
- Carnevale, N.T., and Hines, M.L. (2006). *The NEURON Book* (Cambridge University Press).
- Harris, K.D., and Shepherd, G.M.G. (2015). *Nature Rev. Neurosci.* 18, 170–181.
- Hay, E., Hill, S., Schumann, F., Markram, H., and Segev, I. (2011). *PLoS Comp. Biol.* 7, e1002107.
- Kandel, E.R., Markram, H., Matthews, P.M., Yuste, R., and Koch, C. (2013). *Nat. Rev. Neurosci.* 14, 659–664.
- Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al. (2015). *Cell* 162, 648–661.
- Koch, C. (1999). *Biophysics of Computation* (New York, NY: Oxford University Press).
- Koch, C., and Hepp, K. (2006). *Nature* 440, 611–612.
- Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). *J. Physiol.* 500, 409–440.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). *Science* 305, 793–807.
- Markram, H., Muller, E., Ramswamy, S., Reimann, M.W., Abdellah, M., Sanchez, C.A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., et al. (2015). *Cell* 163, this issue, 456–492.
- Okun, M., Steinmetz, N.A., Cossell, L., Iacaruso, M.F., Ko, H., Barthó, P., Moore, T., Hofer, S.B., Mscic-Flogel, T.D., Carandini, M., and Harris, K.D. (2015). *Nature* 521, 511–515.
- Ramswamy, S., Courcol, J.-D., Abdellah, M., Adaszewski, S., Antille, N., Arsever, S., Guy Antoine, A.K., Bilgili, A., Brukau, Y., Chalimourda, A., et al. (2015). *Front. Neural Circuits* 9, 44.
- Reimann, M.W., Muller, E.B., Ramswamy, S., and Markram, H. (2015). *Front. Comput. Neurosci.* 9, 28.
- Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). *Science* 327, 587–590.
- Strassberg, A.F., and DeFelice, L.J. (1993). *Neural Comput.* 5, 843–855.
- Turing, A.M. (1950). *Mind* 59, 433–460.
- Wilson, H.R., and Cowan, J.D. (1972). *Biophys. J.* 12, 1–24.
- Xu, W., Huang, X., Takagaki, K., and Wu, J.Y. (2007). *Neuron* 55, 119–129.

Small Amounts of Archaic Admixture Provide Big Insights into Human History

Selina Vattathil¹ and Joshua M. Akey^{1,*}

¹Department of Genome Sciences, University of Washington, 3720 15th Avenue Northeast, Box 355065, Seattle, WA 98195-5065, USA

*Correspondence: akey@uw.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.042>

Modern humans overlapped in time and space with other hominins, such as Neanderthals and Denisovans, and limited amounts of hybridization occurred. Here, we review recent work that has identified archaic hominin sequence that survives in modern human genomes and what these genomic excavations reveal about human evolutionary history.

Introduction

As recently as 30,000 years ago anatomically modern humans coexisted with other hominins, such as Neanderthals and Denisovans (Figure 1A). Advances in ancient DNA extraction and sequencing technologies have enabled high-quality Neanderthal (Prüfer et al., 2014) and Denisovan (Meyer et al., 2012) reference genomes to be developed, and in doing so provided the tools to definitively address an often heatedly debated topic in human evolution: did our ancestors interbreed with archaic humans? Analyses leveraging these archaic whole-genome sequences showed that hybridization did indeed happen between our ancestors and Neanderthals (Green et al., 2010; Yang et al., 2012; Sankararaman et al., 2012; Prüfer et al., 2014) and Denisovans (Reich et al., 2010; Reich et al., 2011). In particular, all non-Africans inherited approximately 2% of their genome from Neanderthal ancestors and individuals of Melanesian ancestry inherited approximately 4%–6% of their genome from Denisovan ancestors. Note, the number of hybridization events compatible with the observed levels of archaic ancestry is currently unknown and difficult to estimate precisely as it requires tenuous assumptions about a range of parameters such as historical population sizes and fitness of hybrid offspring.

More recently, the focus of admixture studies has shifted from testing the hypothesis that archaic admixture occurred to identifying the specific sequences that were inherited from archaic ancestors. The two largest studies to date searched for surviving Neanderthal sequences in whole-genome sequence data from individuals of European and East Asian ancestry (Sankararaman et al., 2014; Vernot and Akey, 2014). While these studies used different methodological approaches (reviewed by Racimo et al., 2015), the key population genetics signatures that allow archaic sequences to be identified are unusually high genetic divergence from the human reference genome and high sequence similarity to the Neanderthal reference genome over long (~30–100 kb) stretches of DNA. In total, Vernot and Akey (Vernot and Akey, 2014) identified 600 Mb of introgressed Neanderthal sequence and Sankararaman et al. (Sankararaman et al., 2014) reported 1.1 Gb of introgressed sequence, representing 20% and 35%, respectively, of the Neanderthal genome. Thus, although any given individual carries only a small amount of Neanderthal sequence, cumulatively a substantial amount of

the Neanderthal genome persists in the modern human population and can be recovered by aggregating across hundreds of individuals.

Beyond characterizing the extent of archaic admixture that occurred in East Asians and Europeans, the identification and analysis of surviving Neanderthal sequence in contemporary individuals has revealed a number of fascinating insights about human evolutionary history. Here, we briefly summarize what has been gleaned from studies of archaic hominin admixture, identify existing gaps in knowledge, and suggest fruitful areas of future inquiry.

Widespread Purging of Neanderthal Sequence in Modern Humans

A striking feature of Neanderthal admixture maps is that the amount of surviving Neanderthal sequence varies considerably across the genome (Figure 1B; Sankararaman et al., 2014, Vernot and Akey 2014). Such marked heterogeneity suggests that there may have been fitness consequences to hybridization and some Neanderthal sequences were deleterious in the background of modern human genomes. Consistent with this hypothesis, the frequency of Neanderthal alleles is negatively correlated with inferred functional importance (Sankararaman et al., 2014), and the odds of observing Neanderthal sequence in a region is inversely proportional to the amount of sequence divergence between the modern human and Neanderthal genome (Vernot and Akey 2014). Both of these signatures are expected if introgressed Neanderthal sequence experienced widespread purifying selection in modern humans. Moreover, the X chromosome is significantly depleted of Neanderthal sequence, with an approximate 5-fold reduction in Neanderthal ancestry compared to autosomes (Sankararaman et al., 2014), suggesting reduced fitness in male hybrid offspring (and perhaps male hybrid sterility), a phenomenon observed in many other species (Orr, 1997). Testis-specific genes are significantly enriched in regions depleted of Neanderthal sequence (Sankararaman et al., 2014), further supporting the inference of reduced fitness in male hybrid offspring. It is important to note that the amount of divergence between modern humans and Neanderthals is lower on the X chromosome, which may reduce power to detect introgressed sequence. Nonetheless, although additional work remains to

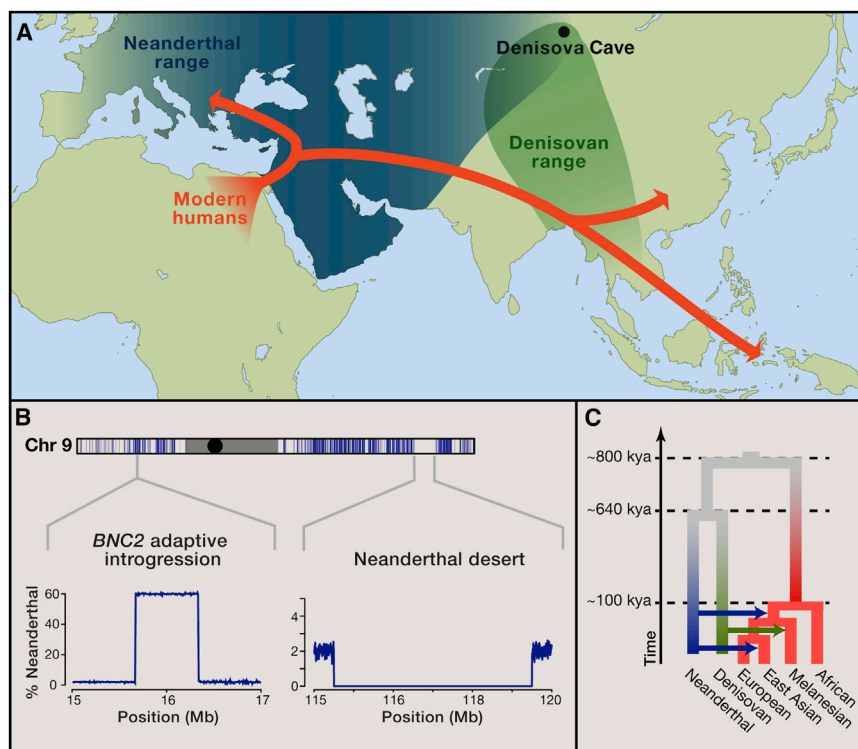


Figure 1. Insights Gleaned from Studying Archaic Hominin Admixture

(A) Schematic illustration of the out of Africa dispersal of modern humans (red). Blue and green shading roughly indicates the regions inhabited by Neanderthals and Denisovans, respectively. Although fossil evidence of Denisovans has been found only at Denisova Cave, the observed pattern of Denisovan admixture suggests they had a wide range that extended southeast.

(B) Distribution of surviving Neanderthal sequence (blue bars) in Europeans and East Asians along chromosome 9 (data from Vernot and Akey 2014). The stylized plots below show the frequency of Neanderthal haplotypes across two genomic regions. An example of putative adaptive introgression (where the Neanderthal haplotype is found at an unusually high frequency) spanning the *BNC2* gene is shown on the left and a ~4 Mb desert of Neanderthal sequence is shown on the right.

(C) Demographic model inferred from patterns of archaic admixture, which includes at least two pulses of Neanderthal admixture and one pulse of Denisovan admixture into individuals of Melanesian ancestry.

they likely had already acquired adaptive mutations that were beneficial to the unique environmental pressures that they evolved in. It is easy to imagine that this variation would have experienced

fully evaluate the inference of male hybrid sterility, it is consistent with the other observed signatures of widespread purifying selection on Neanderthal sequences (Sankararaman et al., 2014).

In addition to these global genomic signatures of purifying selection acting on introgressed Neanderthal sequence, maps of Neanderthal admixture also reveal a number of specific genomic regions that exhibit especially strong depletion of Neanderthal haplotypes (Figure 1B; Sankararaman et al., 2014, Vernot and Akey 2014). These “deserts” of Neanderthal sequence are particularly interesting because they may delineate regions that contribute to uniquely modern human phenotypes. For example, the largest Neanderthal desert on the autosomes spans a ~15 Mb region on chromosome 7 (Vernot and Akey 2014) and contains the gene *FOXP2*, a developmentally important transcription factor that has previously been implicated in speech and language (Enard et al., 2003). However, additional work is necessary to rule out other potential explanations for the origins of Neanderthal deserts, such as inversions (which suppress recombination) or unrecognized extreme demographic histories (such as periods of very small population size in modern humans at the time of admixture). One promising approach in interpreting the biological significance of Neanderthal deserts is the construction of archaic admixture maps for other hominin groups, such as Denisovans. Deserts that replicate across multiple archaic hominins may be of particular biological interest and help refine the genomic substrates of uniquely modern human traits.

Adaptive Introgression

Since archaic hominins lived in Europe and Asia for at least 200,000 years before modern humans migrated to these regions,

strong positive selection in modern humans who acquired it through hybridization (Racimo et al., 2015). Whole-genome Neanderthal admixture mapping has identified several Neanderthal haplotypes that persist at higher frequencies than can be explained by genetic drift (Sankararaman et al., 2014, Vernot and Akey 2014; Racimo et al., 2015). Strikingly, both Sankararaman et al. (2014) and Vernot and Akey (2014) identified three genes in the small set of putatively adaptively introgressed loci that play important roles in hair and skin biology. Although the specific selective pressure underlying the adaptive introgression at these loci is unknown, it is interesting to note that one of these genes, *BNC2* (Figure 1B), has recently been shown to influence skin pigmentation levels in Europeans (Jacobs et al., 2013).

Targeted studies of individual loci have also shown that contact with archaic hominins resulted in adaptive introgression. Perhaps the most dramatic example is that a haplotype of *EPAS1* that was previously shown to confer adaptation to high altitude in Tibetans (likely by lowering hemoglobin concentration) was found to be inherited from a Denisovan-like ancestor (Huerta-Sánchez et al., 2014). Similarly, adaptive introgression of Neanderthal or Denisovan sequences has been hypothesized for HLA (Abi-Rached et al., 2011) and *STAT2* (Mendez et al., 2012), both of which play important roles in the immune response, which is a well-known substrate of selection. Unlike the *EPAS1* haplotype, introgressed haplotypes in these genomic regions are present in multiple human populations.

Thus, archaic admixture appears to have been more than a curious feature of human evolutionary history. Specifically, gene flow from archaic hominins provided important adaptive benefits to our ancestors as they dispersed out of Africa into

new environments. To date, the causal alleles underlying adaptive introgression have not been identified. However, in contrast to recent selective sweeps, which often delimit large genomic regions (~500 kb; see [Akey, 2009](#)), observed adaptively introgressed haplotypes are typically an order of magnitude smaller (median size ~50 kb), which should facilitate fine-scale mapping and experimental studies of putative causal adaptive alleles.

Archaic Admixture Informs Demographic Models of Human History

Initial studies of Neanderthal admixture in geographically diverse populations suggested that all non-African populations had approximately the same levels of Neanderthal ancestry ([Green et al., 2010](#)). Thus, the most parsimonious model to explain these data was that hybridization occurred once in the ancestral population to all present day non-Africans, likely in the Levant shortly after the dispersal of modern humans out of Africa ([Sankararaman et al., 2012](#)). The identification of specific surviving Neanderthal sequence, however, allowed more refined estimates of Neanderthal ancestry in different populations and, unexpectedly, East Asian individuals were found to have on average 20% more Neanderthal sequence than European individuals ([Wall et al., 2013](#), [Sankararaman et al., 2014](#), [Vernot and Akey 2014](#)).

The observation of higher Neanderthal ancestry in East Asians prompted reconsideration of the single-pulse model, and new models suggest at least two distinct admixture events—an initial pulse of admixture into the common ancestor of all present day non-African populations and an additional pulse of admixture into the ancestors of East Asians after their divergence from European populations ([Figure 1C](#); [Vernot and Akey 2014](#)). An alternative explanation that has been proposed is a single pulse of admixture followed by less efficient purging of deleterious Neanderthal sequence in East Asians ([Sankararaman et al., 2014](#)), given their smaller effective population sizes ([Keinan et al., 2007](#)). More recent analyses ([Vernot and Akey, 2015](#), [Kim and Lohmueller, 2015](#)) using distinct statistical approaches found that the two-pulse model better explained the data. Even more directly, the recent observation of Neanderthal admixture in a ~40,000 year old early modern human individual from Romania ([Fu et al., 2015](#)) dramatically shows that admixture occurred in multiple places and times. It is important to stress that even a two-pulse model of admixture with Neanderthals is likely a simplification, and as shown by [Vernot and Akey \(2015\)](#) additional admixture scenarios are also compatible with currently available data.

Currently, much less is known about admixture dynamics between modern humans and Denisovans, whose genetic contribution has been definitively found only in Melanesians ([Reich et al., 2010](#); [Reich et al., 2011](#)), with potentially very low levels of Denisovan ancestry in East Asians ([Skoglund and Jakobsson, 2011](#)). Papua New Guineans carry the highest level of Denisovan ancestry (4%–6% of the genome) among all populations studied to date, with other Southeast Asian island populations generally showing Denisovan ancestry levels consistent with indirect acquisition of introgressed sequence through recent gene flow from populations related to Papua New Guineans ([Reich et al., 2011](#)). These findings, combined with the observed frequencies

of Neanderthal ancestry, suggest that modern humans peopled Asia in at least two distinct waves, with one wave taking a southern route and acquiring Denisovan ancestry and a separate wave responsible for colonization of East Asia and Indonesia. The fossil evidence of Denisovans is limited to a single specimen discovered much farther north, in Siberia, but the geographic pattern of Denisovan ancestry strongly suggests that Denisovans had a range that extended into Southeast Asia ([Figure 1A](#)). These inferences demonstrate how studies of archaic admixture can be incorporated with other information to refine demographic models of modern human history.

Future Challenges and Opportunities

Recent studies of archaic hominin admixture in modern humans have provided fascinating new insights into human evolutionary history and the role that hybridization has played in shaping extant patterns of human genetic variation. Despite this progress, considerable gaps in knowledge remain. Perhaps most obviously, it will be of great interest to comprehensively identify Neanderthal and Denisovan sequences in geographically diverse populations. Such data will be critical for refining admixture models and may lead to the identification of additional examples of adaptive introgression. Furthermore, creating maps of surviving archaic sequences across globally diverse populations will enable introgression deserts to be identified with more precision and confidence, which ultimately may allow the genetic substrates of uniquely modern human traits to be discovered.

Another key opportunity is to systematically explore the phenotypic legacy of archaic hominin admixture. Initial analyses have already shown that genetic variation inherited from Neanderthals contributes to a wide range of human diseases ([Sankararaman et al., 2014](#); [Williams et al., 2014](#)). For example, a type 2 diabetes risk haplotype that is common in Latin Americans and East Asians but rare in Europeans has been found to originate in Neanderthals ([Williams et al., 2014](#)). Moreover, the resources to perform genome-wide association studies by integrating electronic medical records with variants that tag archaic haplotypes are becoming increasingly available ([Denny et al., 2013](#)). Such analyses will allow comprehensive insights into how hybridization has contributed to human phenotypic diversity, and whether particular diseases have a disproportionately high burden attributable to risk variants inherited from archaic ancestors.

Furthermore, new methodological tools need to be developed to fully exploit the information contained in catalogs of introgressed sequence identified in geographically diverse populations. In contrast to sequencing ancient DNA from a single fossil, the recovery of surviving archaic sequences in modern humans provides population level data in that the sequences recovered come from multiple archaic ancestors. Thus, in theory, more powerful population genetics inferences are possible compared to cases where genetic information exists for a single individual. Methodological innovations will bring us closer to answering questions such as the size and genetic structure of archaic populations, the number of archaic ancestors that contributed to surviving archaic sequence in modern humans, dates of introgression, and whether sex-biased admixture occurred.

Finally, we anticipate that future studies of archaic admixture will increasingly focus on identifying and characterizing DNA sequences in the genomes of modern humans that were inherited from hominins other than Neanderthals or Denisovans. The study of African populations may be particularly fruitful, as many hominin lineages existed in Africa. Indeed, preliminary genetic evidence of introgression from an unknown hominin ancestor has been observed in several African populations (Hammer et al., 2011; Lachance et al., 2012). The absence of reference genome sequences will complicate the search for gene flow from unknown hominin lineages, although population genetics tools do exist for identifying putatively introgressed sequence without explicitly using an archaic reference genome (Hammer et al., 2011, Vernot and Akey 2014). Despite these challenges, excavating extinct hominin lineages that persist in the DNA of contemporary individuals may be the only way to identify such groups where no ancient DNA or fossil data exists.

In summary, although the amount of surviving archaic sequence in any individual modern human genome is small, archaic admixture has had a profound lasting impact, not only in shaping who we are today but also in informing us about our history. Ironically, studying the scattered remnants of surviving archaic DNA that persist in modern humans may lead to a deeper understanding of why we survived to expand and thrive while our close relatives did not.

REFERENCES

- Abi-Rached, L., Jobin, M.J., Kulkarni, S., McWhinnie, A., Dalva, K., Gragert, L., Babrzadeh, F., Gharizadeh, B., Luo, M., Plummer, F.A., et al. (2011). *Science* 334, 89–94.
- Akey, J.M. (2009). *Genome Res.* 19, 711–722.
- Denny, J.C., Bastarache, L., Ritchie, M.D., Carroll, R.J., Zink, R., Mosley, J.D., Field, J.R., Pulley, J.M., Ramirez, A.H., Bowton, E., et al. (2013). *Nat. Biotechnol.* 31, 1102–1110.
- Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P., and Paabo, S. (2003). *Folia Primatol. (Basel)* 74, 190–191.
- Fu, Q., Hajdinjak, M., Moldovan, O.T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N., Rohland, N., Lazaridis, I., Nickel, B., et al. (2015). *Nature* 524, 216–219.
- Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). *Science* 328, 710–722.
- Hammer, M.F., Woerner, A.E., Mendez, F.L., Watkins, J.C., and Wall, J.D. (2011). *Proc. Natl. Acad. Sci. USA* 108, 15123–15128.
- Huerta-Sánchez, E., Jin, X., Asan, B., Bianba, Z., Peter, B.M., Vinckenbosch, N., Liang, Y., Yi, X., He, M., Somel, M., et al. (2014). *Nature* 512, 194–197.
- Jacobs, L.C., Wollstein, A., Lao, O., Hofman, A., Klaver, C.C., Uitterlinden, A.G., Nijsten, T., Kayser, M., and Liu, F. (2013). *Hum. Genet.* 132, 147–158.
- Keinan, A., Mullikin, J.C., Patterson, N., and Reich, D. (2007). *Nat. Genet.* 39, 1251–1255.
- Kim, B.Y., and Lohmueller, K.E. (2015). *Am. J. Hum. Genet.* 96, 454–461.
- Lachance, J., Vernot, B., Elbers, C.C., Ferwerda, B., Froment, A., Bodo, J.-M., Lema, G., Fu, W., Nyambo, T.B., Rebbeck, T.R., et al. (2012). *Cell* 150, 457–469.
- Mendez, F.L., Watkins, J.C., and Hammer, M.F. (2012). *Am. J. Hum. Genet.* 91, 265–274.
- Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). *Science* 338, 222–226.
- Orr, H.A. (1997). *Annu. Rev. Ecol. Syst.* 28, 195–218.
- Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., et al. (2014). *Nature* 505, 43–49.
- Racimo, F., Sankararaman, S., Nielsen, R., and Huerta-Sánchez, E. (2015). *Nat. Rev. Genet.* 16, 359–371.
- Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L., et al. (2010). *Nature* 468, 1053–1060.
- Reich, D., Patterson, N., Kircher, M., Delfin, F., Nandineni, M.R., Pugach, I., Ko, A.M.-S., Ko, Y.-C., Jinam, T.A., Phipps, M.E., et al. (2011). *Am. J. Hum. Genet.* 89, 516–528.
- Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., Patterson, N., and Reich, D. (2014). *Nature* 507, 354–357.
- Sankararaman, S., Patterson, N., Li, H., Pääbo, S., and Reich, D. (2012). *PLoS Genet.* 8, e1002947.
- Skoglund, P., and Jakobsson, M. (2011). *Proc. Natl. Acad. Sci. USA* 108, 18301–18306.
- Vernot, B., and Akey, J.M. (2014). *Science* 343, 1017–1021.
- Vernot, B., and Akey, J.M. (2015). *Am. J. Hum. Genet.* 96, 448–453.
- Wall, J.D., Yang, M.A., Jay, F., Kim, S.K., Durand, E.Y., Stevison, L.S., Gignoux, C., Woerner, A., Hammer, M.F., and Slatkin, M. (2013). *Genetics* 194, 199–209.
- Williams, A.L., Jacobs, S.B., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., García-Ortiz, H., Gómez-Vázquez, M.J., Burt, N.P., Aguilar-Salinas, C.A., et al.; SIGMA Type 2 Diabetes Consortium (2014). *Nature* 506, 97–101.
- Yang, M.A., Malaspina, A.-S., Durand, E.Y., and Slatkin, M. (2012). *Mol. Biol. Evol.* 29, 2987–2995.

Beyond Molecular Codes: Simple Rules to Wire Complex Brains

Bassem A. Hassan^{1,2,*} and P. Robin Hiesinger^{3,4,*}

¹Center for the Biology of Disease, VIB, 3000 Leuven, Belgium

²Center for Human Genetics, University of Leuven School of Medicine, 3000 Leuven, Belgium

³Division of Neurobiology, Institute for Biology, Freie Universität Berlin, 14195 Berlin, Germany

⁴NeuroCure Cluster of Excellence, Charité Universitätsmedizin Berlin, 10117 Berlin, Germany

*Correspondence: bh@kuleuven.be (B.A.H.), robin.hiesinger@fu-berlin.de (P.R.H.)

<http://dx.doi.org/10.1016/j.cell.2015.09.031>

Molecular codes, like postal zip codes, are generally considered a robust way to ensure the specificity of neuronal target selection. However, a code capable of unambiguously generating complex neural circuits is difficult to conceive. Here, we re-examine the notion of molecular codes in the light of developmental algorithms. We explore how molecules and mechanisms that have been considered part of a code may alternatively implement simple pattern formation rules sufficient to ensure wiring specificity in neural circuits. This analysis delineates a pattern-based framework for circuit construction that may contribute to our understanding of brain wiring.

Introduction

The brain, as we neuroscientists like to say, is *really* complex. A good deal of our efforts are therefore dedicated to figuring out just how this apparent complexity is generated: where does the information to build a brain come from, and how is such information turned into synapse-specific wiring? We call this the “brain wiring problem.”

We know how it is not done. For example, there cannot be a genetic blueprint that describes every synaptic connection in the way a blueprint of a microchip or electrical wiring diagram does (Figure 1A). Why is this? A blueprint can clearly be envisioned that precisely matches any given neural circuit, allowing for it to be reproducibly built. However, the complexity of such a blueprint exactly equals the complexity of the actual wiring diagram. In other words, this solution generates a new problem that is just as difficult: How is the blueprint generated? It is like answering the question of how life evolved on earth by arguing that it may well have arrived here from a different planet. This is indeed a solution, but an unsatisfying one because it leaves the equally difficult and interesting question of how life evolved on some other planet unresolved.

In this Perspective, we will discuss the concept of developmental algorithms as a solution to the problem of how only a few thousand genes can produce complexity in brain wiring. We will begin by defining some of the key terminology. Next, we consider the evidence for deterministic molecular codes that could define connections, akin to blueprints defining electrical circuits. We will then discuss examples of molecules that were once envisioned to be part of a code but were subsequently revealed to execute developmental rules that integrate stochastic processes in the development of neural circuits. These examples highlight how complicated structures can be generated through simple pattern formation rules rather than elaborately defined addresses. In the last section, we will discuss the difference between an understanding of brain wiring in terms of devel-

opmental rules versus the more common focus on mechanisms of individual molecules that execute those rules.

Genetically Encoded Stochastic Invariability?

In order to understand to what extent and how genes can encode brain wiring, we need to first establish a few simple definitions. First, we define developmental outcomes as “genetically encoded” if environmental input does not contribute any instructive information to that outcome. A common assumption is that a genetically encoded process is invariable, but this would be a misunderstanding. For example, grafting two branches from the same apple tree on two different root stocks will generate two genetically identical trees with non-identical branching patterns—just as no two dendritic arbors of genetically identical Purkinje cells ever look exactly the same (Figure 1B). A simple algorithm can generate an invariable branching pattern (Prusinkiewicz and Lindenmayer, 1990) (Figure 1C); however, introduction of random inaccuracies can create variability in such a system (Figure 1D). Small environmental differences undoubtedly play a role in generating the differences observed in branching patterns between genetically identical specimens. However, this type of differential environmental input does not contain information for any specific branching pattern; rather, the developmental program ensures robustness of functionally important aspects and allows for variability otherwise. By our definition, such a process is genetically encoded, because the environmental input did not contribute any instructive information to generate that outcome: in both the apple tree and the Purkinje cell, the precise position of branches is irrelevant as long as the final branched structure covers a certain region in a specific manner. Hence, genetically encoded developmental algorithms can produce similar but non-identical structures in the brain.

Small environmental differences are a random variable and therefore define a stochastic process that can lead to variability in the outcome. Conversely, a system is defined as deterministic

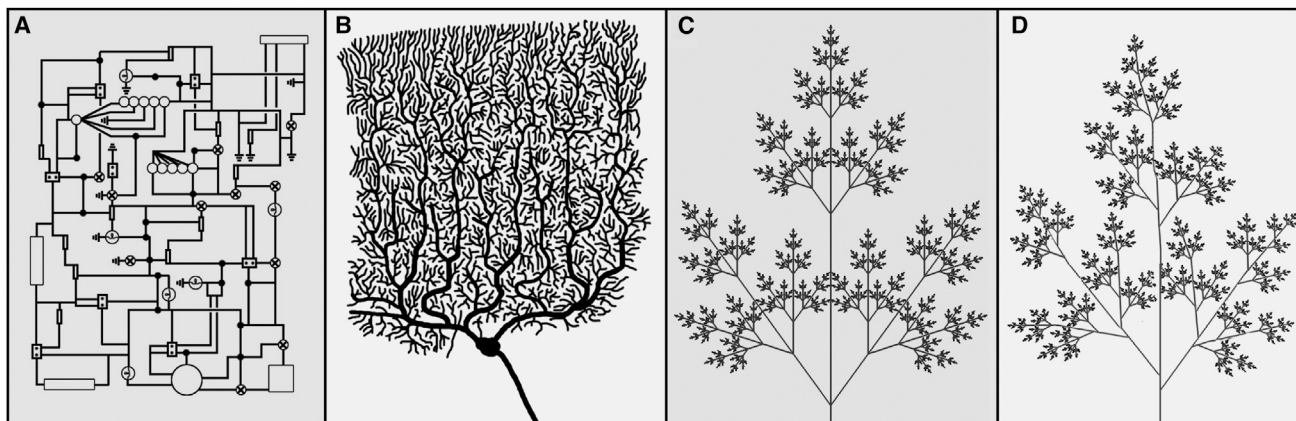


Figure 1. From Deterministic Blueprints to Stochastically Branched Structures in Biology

(A) Schematic of a hypothetical electrical blueprint with deterministic definition of all contacts.

(B) Schematic drawing of a Purkinje cell after a well-known drawing from Ramón y Cajal. The precise branching pattern, number, and placement of dendritic endings is variable.

(C) A simple computer-generated branched structure. The deterministic definition of all branches is generated by a few lines of code (Lindenmayer system) using L-Studio 4.2.13 by Przemysław Prusinkiewicz and Radek Karwowski.

(D) The same branched structure as in C, but with stochastic pattern changes.

if no randomness is involved in the development of future states of the system. However, invariability in the outcome does not require a deterministic system. As we shall see throughout this Perspective, even when the outcomes appear invariable, the processes that generate them are often stochastic. All genetically encoded developmental signaling events contain stochastic processes. Notch signaling provides an excellent example. This pathway encodes a molecular mechanism that breaks the symmetry in cell differentiation by ensuring that only one daughter cell becomes cell type A, the other type B. The outcome of exactly one cell type A and B is invariable, but it is crucial to consider that the fate from the perspective of the individual cell is in fact stochastic, because it is impossible to predict which of the two cells will be A and which will be B. We call this a genetically encoded stochastic process with an invariable, or highly stereotyped, outcome. Thus, be it during the development of the vertebrate heart (de la Pompa and Epstein, 2012) or the fly eye (Carthew, 2007), stochastic processes are parts of genetically encoded developmental programs that lead to highly stereotyped and robust outcomes.

A particularly insightful example of a developmental process critical for brain wiring is the idea that “cells that fire together, wire together” in the mammalian visual system (Shatz, 1996). This process is based on spontaneous activity waves that Shatz and colleagues first saw sweeping over the ferret retina even before these cells are capable of receiving environmental input. While the activity waves are stochastic, they result in stereotyped layer formation in the lateral geniculate nucleus. Some variability occurs (e.g., the precise size of the layers), but no environmental input contributes instructive information to generate the developmental outcome. Hence, we consider this process a genetically encoded stochastic process with a highly stereotyped outcome. As an aside, it follows from these considerations that an “activity-dependent” process can be part of a genetically encoded program. In this Perspective, we only explore such

genetically encoded processes, while the important roles of environmental input in activity-dependent synaptic fine tuning are reviewed elsewhere (Ganguly and Poo, 2013; West and Greenberg, 2011). Obviously, genes do not encode stochastic spontaneous neuronal activity just as they do not encode stochastic branching patterns. Instead, in both cases, gene activity defines the developmental algorithms that lead to such cellular behaviors.

An important lesson from these examples is the use of stochastic processes as an integral and necessary part of developmental algorithms, which contrasts with the view of noise as something that development just has to cope with or minimize in order to create a robust outcome (Clarke, 2012; Melé et al., 2015). Recent work in several fields has highlighted the importance of understanding both stochastic processes and heterogeneity of cellular behavior. These fundamental features of all biological systems are lost when we focus on studying averages (Altschuler and Wu, 2010; Losick and Desplan, 2008). How developmental algorithms generate apparent complexity in brain wiring therefore requires insights into the developmental process that are often non-intuitive and quantitative.

Molecular Codes and the “Complexity Reduction Model”

What fundamental solutions to the brain wiring problem do molecular codes offer? Genes encode molecules, and molecules can theoretically provide combinatorial codes of almost any complexity. The success story of molecular biology and gene discovery provides us with ample examples. The obvious candidates for establishing molecular codes in intercellular, synapse-specific interactions are secreted and membrane-associated “guidance cues,” their receptors, and cell adhesion molecules (Kolodkin and Tessier-Lavigne, 2011). These cues include molecules that belong to “canonical” guidance cue families (Netrins, Slits, Semaphorins, and Ephrins) as well as cell adhesion molecules of the immunoglobulin or cadherin superfamilies. The

canonical secreted and membrane-associated guidance cues function at long range or short range to mediate attractive or repulsive signals; cell adhesion molecules may function through direct contact-mediated homophilic or heterophilic interactions. These molecules and mechanisms are reviewed in detail elsewhere (Kolodkin and Tessier-Lavigne, 2011; Raper and Mason, 2010; Yogeve and Shen, 2014).

Many of these genes can produce mRNAs resulting from different splice variants or utilization of multiple promoters, further increasing the numerical potential for different combinatorial codes. The most impressive examples are invertebrate Dscams and vertebrate Protocadherins, since members of each family are present in thousands of different splice variants that are required for wiring specificity (Lefebvre et al., 2012; Schmucker et al., 2000; Zipursky and Sanes, 2010). In addition, many of the key growth factor and embryonic patterning pathways such as Wnt, FGF, EGF, and BMP have been found to be required for brain wiring (Charron and Tessier-Lavigne, 2005; Srahn et al., 2006). Thus, a numerically large array of molecular combinations is, in principle, available for neurons to both display and respond to.

A common feature of guidance cues and receptors is their cell-specific expression and spatiotemporally dynamic localization (Chan et al., 2011; Williamson et al., 2010; Zschätzsch et al., 2014). This observation further supports their potential roles in establishing molecular codes; temporal coding using the same guidance receptor can generate a “code in time” to specify target areas (Petrovic and Hummel, 2008; Yogeve and Shen, 2014). Hence, a picture emerges in which different cells may express distinct combinations of guidance cues and receptors; different combinations of these molecules can then be presented at distinct places and at specific times during brain wiring to provide unique targeting and synapse formation signals. It is also often argued that early connectivity events take place in much less complicated wiring environments than the final pattern might indicate. The resulting model thus assumes a stepwise process successively restricting possible targeting choices such that the problem of choosing among thousands of options may never occur. We call this the “complexity reduction model.”

Numerous elegant mechanisms that contribute to complexity reduction have been put forth. For example, classic guidance cues can form gradients that help in the parallel targeting of many axons simultaneously (Kolodkin and Tessier-Lavigne, 2011; Yogeve and Shen, 2014). Parallelized targeting creates repetition and thus redundancy in information encoding: a molecularly encoded blueprint for such a system only needs to specify one of the repetitive units. In addition, even a limited number of guidance cues may be consistent with the apparent complexity of brain wiring if specificity is achieved through regulation of guidance cues in space and time, through protein expression levels, or through the utilization of heteromultimeric receptor machinery. The complexity reduction model is thereby proposed to provide a solution to the brain wiring problem: we may not yet have worked out all the details of when, where, and what kinds of combinatorial molecular codes occur, but there seems to be no fundamental problem. Or is there?

The Molecular Code and Its Discontents

After establishing the idea of the complexity reduction model, we must ask how the underlying stepwise, spatiotemporal code is generated. How do specific combinations of guidance cues selectively and precisely get to be at the right time and place to function as meaningful synapse specification signals? Here again, we face the blueprint problem: the establishment of a deterministic, spatiotemporally precise molecular code would be a problem of comparable complexity to the wiring diagram that it is supposed to explain.

The blueprint problem holds irrespective of whether the code is determined through numbers of molecular cues, temporal control of molecular cues, or differential expression levels of molecular cues; it is not obviously easier to control the precise proteins levels of one cue in space and time than to control the precise combination of several cues only in space (Chan et al., 2011). The example of gradients highlights one way to create repetition of similar structures and thereby reduce the amount of information, or codes, needed. Complexity can thus theoretically be reduced through the exact repetition of a precise and deterministic address code system. However, repetitive structures in brain wiring do typically, and maybe without exception, allow for some level of variability, revealing an underlying stochastic process. This variability may be functionally irrelevant (like differences in precise branching patterns, comp. Figures 1B–1D), but it can reveal the underlying developmental rules and mechanisms: molecular cues that provide approximate guidance for many axons in parallel do not specify a precise one-to-one address code. Hence, variability may not just be due to the repetition of a slightly imprecise address code but, rather, an inherent outcome of a stochastic process that neither requires nor generates an address code.

The complexity reduction model allows for variability by making guidance cues less precise. But how does lack of precision ensure robustness of the developmental process? We have already argued that noise is not simply an artifact that biology has to “live with” and try to reduce but is an integral part of how a developmental algorithm functions in brain wiring. How do guidance molecules deal with or even utilize noise? Numerous influential studies on cell adhesion and guidance receptor functions in different model systems provide us with ample examples for a new and surprising understanding of what these molecules do. For example, we now know that, in *Drosophila*, an individual neuron’s choice of one out of thousands of Dscam1 isoforms is indeed unpredictable (Miura et al., 2013). Interestingly, a wealth of groundbreaking work on Dscam1 has revealed a function for this non-deterministic isoform choice that is quite different from a precise address code: both Dscam and Protocadherins in vertebrates serve a primary role in mediating self-avoidance, which requires distinction between self and non-self cell surfaces (Lefebvre et al., 2012; Zipursky and Sanes, 2010). Similarly, cell-intrinsic stochastic recycling of EGFR in the growth cones of higher-order visual system neurons in *Drosophila* is required to form a highly stereotyped axonal branching pattern (Zschätzsch et al., 2014). Variability does not arise in these examples by making a molecular code less precise. Instead, variability is a necessary outcome of pattern formation processes following simple rules. Specifically,

Dscams and Protocadherins implement a simple pattern formation rule without providing a “cue” for axon targeting or synapse specification (Kise and Schmucker, 2013; Zipursky and Sanes, 2010). Large numbers of randomly chosen isoforms serve this function, similar to randomly coded remote garage door openers; as long as there are enough different isoforms, the likelihood is sufficiently low that your neighbor’s system has the same random recognition code as you do. This beautiful role of the many isoforms of Dscams and Protocadherins is more akin to a pattern formation process than a molecular synapse specification code.

The idea of “non-cue” functions of guidance molecules is further highlighted by the identification of cell-intrinsic functions. If no target is involved in an intrinsic axonal targeting or branching choice, then the implicated molecules cannot function as external cues or address code (Petrovic and Schmucker, 2015). Examples include neuropilin-1 and semaphorin-3a in the mammalian olfactory system (Imai et al., 2009), the protocadherin Flamingo in the fly visual system (Schwabe et al., 2013), and, indeed, Dscam1 in mechanosensory neurons (He et al., 2014). Variability in the outcome arises not from an imprecision of a molecular code but because these processes are intrinsically noisy and because they generate patterns, rather than specify connections.

A picture emerges of how developmental algorithms can “encode” synaptic specificity in neural circuits. The question thereby is: what exactly are these developmental rules, and how can they explain wiring specificity in the nervous system?

The Simple Rules that Can

Any comprehensive mechanistic explanation for brain wiring must explain three fundamental characteristics in a single conceptual framework: wiring specificity, wiring variability, and the inclusion of stochastic processes during development. A rules-based framework does just that. Variability in neuronal branching patterns and neural circuit architecture offers a first glimpse into underlying developmental rules that generate complicated connections. For example, a set of simple rules that includes spacing between axons and self-avoidance can explain both the robust and variable properties of axonal targeting patterns. A classic and fundamental simple pattern formation rule is lateral inhibition, which is often molecularly implemented through Notch signaling. Indeed, recent work in *Drosophila* shows that this classical pattern formation principle in cell differentiation is actively employed by axons of postmitotic neurons during neural circuit assembly the *Drosophila* brain (Langen et al., 2013). Here, stochastic patterning leads to spacing with a defined average between neighboring axons without specifying either the precise number or distance between them. This is an example of a simple rule that does not by itself specify a synaptic connection but must function as part of a larger developmental algorithm.

In another recent example, the seemingly complex wiring principle of the *Drosophila* visual system known as “neural superposition” was computationally modeled based on three simple rules (Langen et al., 2015). Rather than explaining the simultaneous targeting of ~5,000 growth cones through mechanisms of attractive and repulsive molecular cues, the entire synaptic specification process can be explained through pattern formation principles executed as a concatenation of simple genetically

encoded subprograms (Hiesinger et al., 2006; Langen et al., 2015). This example presents all features discussed here: wiring specificity (almost error-free connectivity), variability in the precise spatial placement of pre- and postsynaptic elements (which may or may not be functionally relevant), and stochastic processes throughout brain development from neuronal cell fate choice to growth cone dynamics. Importantly, previously identified roles of cell adhesion molecules in this process fit seamlessly into this framework as implementers of pattern formation rules, rather than molecular address codes.

An important aspect of brain wiring models based on simple pattern formation rules is the “sorting together” of presumptive synaptic partners, such that synapse formation is likely to occur between meaningful partners (Hiesinger et al., 2006). Indeed, it is a curious fact that the actual process of synapse formation appears to be astonishingly non-specific across species. Neurons that innervate incorrect target regions generally will form synapses wherever they end up, however wrong the targets. In fact, if given no other choice, neurons readily form synapses with themselves (so-called autapses) that are functionally indistinguishable from synapses in the brain (Bekkers and Stevens, 1991). Hence, as an ingredient for synapse-specific brain wiring, non-specific synapse formation only makes sense in the context of a larger developmental algorithm. Specifically, promiscuous synapse formation can be guided by precise sorting of the right partners; this is observed, for example, in the mammalian olfactory system or the fly visual system (Hiesinger et al., 2006; Imai et al., 2009). Alternatively, activity-dependent pruning can function in sculpting specific layers harboring specific synaptic connections, as is observed in the lateral geniculate nucleus of the vertebrate visual system (Shatz, 1996). In addition, developmental rules inconsistent with address codes are revealed by human patient data. Axons from the left eye normally establish connection in the visual areas of the right brain hemisphere and vice versa. In 2009, physicians described the case of a child born without a right brain hemisphere who nonetheless developed both left and right visual fields, in both the lateral geniculate nucleus and visual cortex, in the left hemisphere alone, resulting in near normal vision (Muckli et al., 2009). Taken together, a picture emerges of how a purely genetically encoded developmental program can lead to synaptic specificity in a neural circuit through a concatenation of simple genetically encoded subprograms that employ stochastic processes and allow for flexibility, without the need for an address code.

What simple rules can, together with unspecific synapse formation, create synapse specific wiring? Rules can pattern axonal and dendritic architectures that result in specific synaptic partners. For example, the development of the complicated dendritic tree of a Purkinje cell (Figure 1B) has been modeled and computer simulated using simple dynamic processes, including stochastic terminal branching and retraction triggered by dendritic contact (Fujishima et al., 2012). Based on similar “on-contact” rules, we would like to propose a simple theoretical experiment. This experiment reveals how two simple pattern formation rules can determine synaptic specificity when assuming that synaptogenesis can occur between any presynaptic-postsynaptic contact. We envision a schematic brain structure (Figure 2) organized into layers and columns with dynamically extending

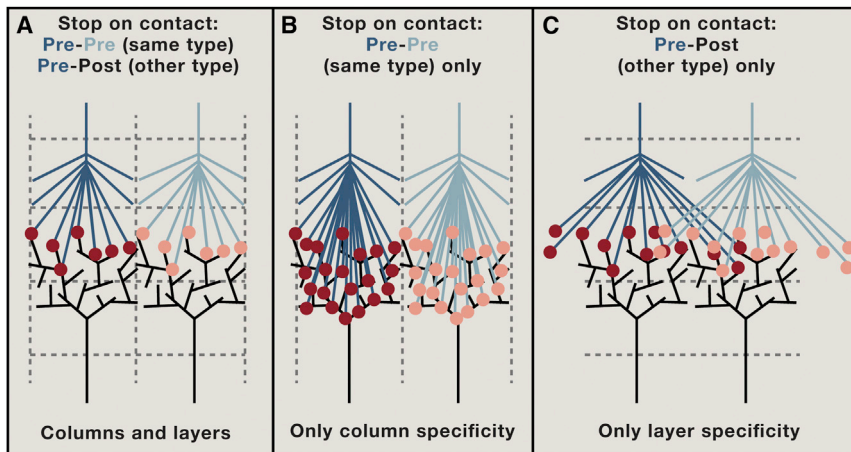


Figure 2. A Theoretical Experiment: Given Promiscuous Synaptogenesis at Any Contact Site, Two Simple Rules Are Sufficient to Generate Layer- and/or Column-Specific Synaptic Contacts

(A) The rule “stop on pre-pre (or same cell type) contact” prevents overlap of neighboring, parallel presynaptic terminals, leading to tiling in columns. The rule “stop on pre-post (or other cell type) contact” prevents overlap within the column; the area where pre- and postsynaptic terminals meet defines a layer. Synapses can subsequently form “unspecifically” between any pre-post contact and are yet restricted to a specific column and layer. (B) The “pre-pre” rule is sufficient to maintain columns, but without a “pre-post” rule, overlap between different cell types lead to loss of a restricted layer. (C) The “pre-post” rule is sufficient to maintain layers, but without a “pre-pre” rule, overlap between the same presynaptic cell types leads to loss of columnar restriction.

presynaptic axon terminals (blue) and postsynaptic dendritic trees (black). Synapses (red/pink dots) will form promiscuously at pre-postsynaptic contact sites. In this minimal setup, the two simple rules “stop on presynaptic-presynaptic contact” and “stop on presynaptic-postsynaptic contact” can generate both layer and column specificity (Figure 2A). Specifically, stop on contact between presynaptic axon terminals and dendritic branches restricts the area of synapse formation to a specific layer; stop on contact between presynaptic axon terminals ensures that neither invade each other’s column. Correspondingly, loss of the “stop on pre-post contact” rule reduces layer specificity while preserving columns (Figure 2B); loss of the “stop on pre-pre contact” rule reduces column specificity while preserving layers (Figure 2C). A similar observation has recently been made for starburst amacrine cells, which form autapses when self-avoidance is perturbed and fail to form connections with other cells when their contact recognition is perturbed. (Kostadinov and Sanes, 2015). Hence, simple rules such as those executed by cell adhesion molecules previously interpreted as ‘guidance cues’ or ‘recognition codes’ can easily generate variety in columnar and layer organization and synaptic specificity. Local interaction rules can be iteratively applied across layers and columns and thereby provide complexity reduction. Finally, the same rule may be executed by different molecules in different systems, as seen for Dscams and protocadherins.

The examples discussed here showcase how simple rules can explain the establishment of synaptic specificity in seemingly complex wiring diagrams to a significant extent. All of these examples are based on iteration of simple rules and thereby create repetitive structures. We speculate that less-repetitive organization can result in brain regions where different developmental algorithms overlap. However, many examples remain where more deterministic solutions based on true guidance and/or matching cues appear to provide a satisfactory explanation. For example, specific laminae in the vertebrate retina are thought to be defined by distinct guidance receptors (Matsuoka et al., 2011). In the *Drosophila* olfactory system, the Teneurins *ten-m* and *ten-a* are proposed to function as homophilic “match-making” molecules (Hong et al., 2012). Importantly, the idea of match-making poses some significant constraints on stochasticity, as both pre- and

postsynaptic sides must have “matching” molecular partners and thus deterministic molecular recognition pairs. As a general mechanism for synaptic specification throughout circuit assembly, the idea of match-making is not easily reconciled with stochastic developmental processes. On the other hand, the match-making roles found for these two proteins occur very late in the developmental process to distinguish between few targeting choices prior to synapse formation; thereby, match-making can be understood as an elegant terminal subprogram of a larger developmental algorithm that leads to synaptic specificity.

On the Relation between Developmental Rules and Molecular Mechanisms

Much insight into neural circuit assembly has been gained from single-mutant gene studies that disrupt development. Such experiments are often designed to reveal molecular mechanisms, including attractive or repulsive interactions requiring cell surface receptors during neural circuit assembly. The perspective of developmental rules differs from this approach in the following way: a molecular mechanism executes a developmental rule but may not reveal the rule itself. In contrast, the developmental algorithm is defined as the set of rules that are sufficient to generate robust and precise wiring. Developmental rules can be formulated independent of the molecular mechanisms that execute them, as shown in the theoretical experiment above (Figure 2). More specifically, classical molecular mechanisms of guidance cues and receptors include homophilic and heterophilic binding, both of which can implement either attractive or repulsive responses. But neither of these mechanisms by themselves reveal their roles as guidance cues. For example, the molecular mechanism of Dscam1 (and all its isoforms) is homophilic repulsion; however, this mechanism does not reveal its true role in implementing the simple pattern formation rule of self-avoidance. In contrast, the self-avoidance rule can be quantitatively formulated and understood in the absence of molecular knowledge.

How then can we identify developmental rules independent of the molecular mechanisms that execute them? Curiously, the most common route has remained molecular perturbation

experiments. The assumption in any molecular perturbation experiment is that taking a specific part out will reveal meaningful behavior of the system through the observed response and thus define the function of that specific part. The added hope is that the part whose role is revealed tells us something about the developmental rule that it executed. However, the system is likely to exhibit compensatory responses and secondary effects that may be difficult to interpret. Molecular perturbation experiments are therefore more likely to reveal underlying rules when the perturbation is carried out with high spatial and temporal resolution. For example, analyses of N-cadherin and Flamingo in the fly visual system revealed general principles only through detailed investigation of individual mutant growth cones in relation to identified wild-type or mutant neighbors (Schwabe et al., 2013, 2014).

The characterization of developmental rules independent of the molecules that execute them is probably best achieved through live observation with or without spatiotemporally controlled perturbation. However, the live observation approach demands the ability to observe growth cone behavior and synaptogenesis during a relevant time period in a developing neural circuit without interference and with sufficient spatial and temporal resolution. Where this has been achieved, live observation of individual neurons and their interactions over time yielded important insight into the temporal succession, and thus causal constraints, for underlying brain wiring processes (Langen et al., 2015). In another example, recent live imaging of neuronal migration in the zebrafish retina revealed unexpected cellular behavior leading to amacrine cell lamination (Chow et al., 2015). It will therefore be interesting to extend the live observation to other cell types in the vertebrate retina and the fly olfactory system to see whether cellular behaviors are best explained by code-based target selection mechanisms or simple pattern formation rules.

However we attempt to break down the often quoted “daunting complexity” of the brain, a complete solution to the brain wiring problem may bear more similarity to other developmental tissues than our intuition at first suggests. Rules like lateral inhibition, self-avoidance, and gradient-based patterning, as well as underlying mechanisms like heterophilic interaction, homophilic repulsion, and molecular gradients, are well-established facets of the development of all tissues. As such, brain wiring is likely to a large extent an example of particularly complicated developmental patterning rather than a special problem onto itself. While it may seem safest to the engineer in us to explain brain complexity with an equally complicated code, the history of developmental biology teaches us differently again and again: simple rules!

ACKNOWLEDGMENTS

We would like to thank our laboratories and many friends and colleagues for enduring discussions leading to this text. In particular, we want to thank for critical and very valuable discussions and comments on this article: Tom Claudinin, Alex Kolodkin, Dietmar Schmucker, Nevine Shalaby, Steven Altschuler, Lani Wu, and Joris de Wit. Work in the Hassan laboratory is funded by VIB, University of Leuven, FWO (G.0503.12), belspo (WiBrain Network), EMBO, and the European Commission's Marie Skłodowska Curie programs. Work in the Hiesinger laboratory at the Freie Universität Berlin is funded by the NeuroCure

Cluster of Excellence, the NIH (RO1EY018884, RO1EY023333), the Muscular Dystrophy Association, the Deutsche Forschungsgemeinschaft (SFB958), and the FU Berlin.

REFERENCES

- Altschuler, S.J., and Wu, L.F. (2010). Cellular heterogeneity: do differences make a difference? *Cell* 141, 559–563.
- Bekkers, J.M., and Stevens, C.F. (1991). Excitatory and inhibitory autaptic currents in isolated hippocampal neurons maintained in cell culture. *Proc. Natl. Acad. Sci. USA* 88, 7834–7838.
- Carthew, R.W. (2007). Pattern formation in the *Drosophila* eye. *Curr. Opin. Genet. Dev.* 17, 309–313.
- Chan, C.C., Epstein, D., and Hiesinger, P.R. (2011). Intracellular trafficking in *Drosophila* visual system development: a basis for pattern formation through simple mechanisms. *Dev. Neurobiol.* 71, 1227–1245.
- Charron, F., and Tessier-Lavigne, M. (2005). Novel brain wiring functions for classical morphogens: a role as graded positional cues in axon guidance. *Development* 132, 2251–2262.
- Chow, R.W., Almeida, A.D., Randlett, O., Norden, C., and Harris, W.A. (2015). Inhibitory neuron migration and IPL formation in the developing zebrafish retina. *Development* 142, 2665–2677.
- Clarke, P.G. (2012). The limits of brain determinacy. *Proc. Biol. Sci.* 279, 1665–1674.
- de la Pompa, J.L., and Epstein, J.A. (2012). Coordinating tissue interactions: Notch signaling in cardiac development and disease. *Dev. Cell* 22, 244–254.
- Fujishima, K., Horie, R., Mochizuki, A., and Kengaku, M. (2012). Principles of branch dynamics governing shape characteristics of cerebellar Purkinje cell dendrites. *Development* 139, 3442–3455.
- Ganguly, K., and Poo, M.M. (2013). Activity-dependent neural plasticity from bench to bedside. *Neuron* 80, 729–741.
- He, H., Kise, Y., Izadifar, A., Urwyler, O., Ayaz, D., Parthasarathy, A., Yan, B., Erfurth, M.L., Dascenco, D., and Schmucker, D. (2014). Cell-intrinsic requirement of Dscam1 isoform diversity for axon collateral formation. *Science* 344, 1182–1186.
- Hiesinger, P.R., Zhai, R.G., Zhou, Y., Koh, T.W., Mehta, S.Q., Schulze, K.L., Cao, Y., Verstreken, P., Claudinin, T.R., Fischbach, K.F., et al. (2006). Activity-independent prespecification of synaptic partners in the visual map of *Drosophila*. *Curr. Biol.* 16, 1835–1843.
- Hong, W., Mosca, T.J., and Luo, L. (2012). Teneurin instruct synaptic partner matching in an olfactory map. *Nature* 484, 201–207.
- Imai, T., Yamazaki, T., Kobayakawa, R., Kobayakawa, K., Abe, T., Suzuki, M., and Sakano, H. (2009). Pre-target axon sorting establishes the neural map topography. *Science* 325, 585–590.
- Kise, Y., and Schmucker, D. (2013). Role of self-avoidance in neuronal wiring. *Curr. Opin. Neurobiol.* 23, 983–989.
- Kolodkin, A.L., and Tessier-Lavigne, M. (2011). Mechanisms and molecules of neuronal wiring: a primer. *Cold Spring Harb. Perspect. Biol.* 3, 3.
- Kostadinov, D., and Sanes, J.R. (2015). Protocadherin-dependent dendritic self-avoidance regulates neural connectivity and circuit function. *eLife* 4, 4.
- Langen, M., Koch, M., Yan, J., De Geest, N., Erfurth, M.L., Pfeiffer, B.D., Schmucker, D., Moreau, Y., and Hassan, B.A. (2013). Mutual inhibition among postmitotic neurons regulates robustness of brain wiring in *Drosophila*. *eLife* 2, e00337.
- Langen, M., Agi, E., Altschuler, D.J., Wu, L.F., Altschuler, S.J., and Hiesinger, P.R. (2015). The Developmental Rules of Neural Superposition in *Drosophila*. *Cell* 162, 120–133.
- Lefebvre, J.L., Kostadinov, D., Chen, W.V., Maniatis, T., and Sanes, J.R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* 488, 517–521.
- Losick, R., and Desplan, C. (2008). Stochasticity and cell fate. *Science* 320, 65–68.

- Matsuoka, R.L., Nguyen-Ba-Charvet, K.T., Parry, A., Badea, T.C., Chédotal, A., and Kolodkin, A.L. (2011). Transmembrane semaphorin signalling controls laminar stratification in the mammalian retina. *Nature* 470, 259–263.
- Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al.; GTEx Consortium (2015). Human genomics. The human transcriptome across tissues and individuals. *Science* 348, 660–665.
- Miura, S.K., Martins, A., Zhang, K.X., Graveley, B.R., and Zipursky, S.L. (2013). Probabilistic splicing of *Dscam1* establishes identity at the level of single neurons. *Cell* 155, 1166–1177.
- Muckli, L., Naumer, M.J., and Singer, W. (2009). Bilateral visual field maps in a patient with only one hemisphere. *Proc. Natl. Acad. Sci. USA* 106, 13034–13039.
- Petrovic, M., and Hummel, T. (2008). Temporal identity in axonal target layer recognition. *Nature* 456, 800–803.
- Petrovic, M., and Schmucker, D. (2015). Axonal wiring in neural development: Target-independent mechanisms help to establish precision and complexity. *BioEssays* 37, 996–1004.
- Prusinkiewicz, P., and Lindenmayer, A. (1990). *The Algorithmic Beauty of Plants* (Springer).
- Raper, J., and Mason, C. (2010). Cellular strategies of axonal pathfinding. *Cold Spring Harb. Perspect. Biol.* 2, a001933.
- Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). *Drosophila Dscam* is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101, 671–684.
- Schwabe, T., Neuert, H., and Clandinin, T.R. (2013). A network of cadherin-mediated interactions polarizes growth cones to determine targeting specificity. *Cell* 154, 351–364.
- Schwabe, T., Borycz, J.A., Meinertzhagen, I.A., and Clandinin, T.R. (2014). Differential adhesion determines the organization of synaptic fascicles in the *Drosophila* visual system. *Curr. Biol.* 24, 1304–1313.
- Shatz, C.J. (1996). Emergence of order in visual system development. *Proc. Natl. Acad. Sci. USA* 93, 602–608.
- Srahna, M., Leyssen, M., Choi, C.M., Fradkin, L.G., Noordermeer, J.N., and Hassan, B.A. (2006). A signaling network for patterning of neuronal connectivity in the *Drosophila* brain. *PLoS Biol.* 4, e348.
- West, A.E., and Greenberg, M.E. (2011). Neuronal activity-regulated gene transcription in synapse development and cognitive function. *Cold Spring Harb. Perspect. Biol.* 3, 3.
- Williamson, W.R., Yang, T., Terman, J.R., and Hiesinger, P.R. (2010). Guidance receptor degradation is required for neuronal connectivity in the *Drosophila* nervous system. *PLoS Biol.* 8, e1000553.
- Yogev, S., and Shen, K. (2014). Cellular and molecular mechanisms of synaptic specificity. *Annu. Rev. Cell Dev. Biol.* 30, 417–437.
- Zipursky, S.L., and Sanes, J.R. (2010). Chemoaffinity revisited: *dscams*, protocadherins, and neural circuit assembly. *Cell* 143, 343–353.
- Zschätzsch, M., Oliva, C., Langen, M., De Geest, N., Ozel, M.N., Williamson, W.R., Lemon, W.C., Soldano, A., Munck, S., Hiesinger, P.R., et al. (2014). Regulation of branching dynamics by axon-intrinsic asymmetries in Tyrosine Kinase Receptor signaling. *eLife* 3, e01699.

Pausing on Polyribosomes: Make Way for Elongation in Translational Control

Joel D. Richter^{1,*} and Jeff Collier^{2,*}

¹Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA

²Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, OH 44106, USA

*Correspondence: joel.richter@umassmed.edu (J.D.R.), jmc71@case.edu (J.C.)

<http://dx.doi.org/10.1016/j.cell.2015.09.041>

Among the three phases of mRNA translation—initiation, elongation, and termination—initiation has traditionally been considered to be rate limiting and thus the focus of regulation. Emerging evidence, however, demonstrates that control of ribosome translocation (polypeptide elongation) can also be regulatory and indeed exerts a profound influence on development, neurologic disease, and cell stress. The correspondence of mRNA codon usage and the relative abundance of their cognate tRNAs is equally important for mediating the rate of polypeptide elongation. Here, we discuss recent results showing that ribosome pausing is a widely used mechanism for controlling translation and, as a result, biological transitions in health and disease.

Introduction

Since translational control became a distinct field of study, the term “control” for many investigators was synonymous with initiation, the first and most complicated phase of protein synthesis. Initiation includes formation of the 43S pre-initiation complex, its association with the mRNA 5' terminal 7mG cap in coordination with the eIF4F (eIF4A, eIF4G, and eIF4E) complex, scanning of the 40S ribosomal subunit to the initiation AUG codon, and joining of the 60S subunit to form the 80S monosome (Hinnebusch 2014).

Cells generally contain a dearth of the cap-binding factor eIF4E (Mamane et al., 2004), and its interaction with eIF4G and hence its ability to recruit the translational apparatus is widely regulated by different classes of protein factors, including the eIF4E binding proteins (4EBPs) (Richter and Sonenberg, 2005). Additionally, scanning of the 40S ribosomal subunit along the mRNA can be impeded by interacting proteins or secondary structure in the 5'UTR. Given these distinct control points, tradition has dictated that initiation would be rate limiting for protein synthesis. Moreover, it makes intuitive sense that the first step in translation would be the most likely to be regulated. However, emerging evidence indicates that polypeptide elongation (ribosome transit) can also be regulatory and indeed may be critical for controlling early development, neural function, and cancer etiology. Here, we review salient observations pointing to an important role for regulated ribosome translocation in diverse biological contexts.

Translational Elongation at a Glance

Translational initiation involves dozens of individual polypeptides, and given its complexity, it is not surprising that distinct sub-steps can be regulated. In comparison, translational elongation is relatively simple. In concert with elongation factor EF-1/EF-TU and guanosine triphosphate (GTP), an aminoacylated tRNA enters the A site of the ribosome; cognate tRNA-mRNA codon recognition then stimulates GTP hydrolysis and eviction of EF-1/EF-TU from the A site. Concomitantly, the ribosome un-

dergoes a conformational shift, stimulating contact between the 3' ends of the aminoacylated tRNA in the A site and the tRNA bearing the polypeptide chain in the P site. When the two tRNAs shift position (A to P and P to E site), peptide bond formation occurs as the polypeptide is now transferred to the aminoacylated tRNA, extending the protein by one amino acid. A second elongation factor, EF-2/EF-G, then enters the A site, hydrolyzing GTP and resetting the ribosome to a conformation competent to receive the next aminoacylated tRNA in the A site. The process repeats itself over and over again (Figure 1).

Despite the “simplicity” of elongation, regulation can and does occur. Indeed, for decades, we have known that ribosomes stall after reading only the first 5–30 codons of mRNAs encoding secreted proteins (Siegel and Walter 1988; Halic et al., 2004). This activity requires the signal recognition particle (SRP), which binds the N terminus of the nascent polypeptide and simultaneously inserts itself into the ribosome A site (Halic et al., 2004). Docking of SRP in the ribosomal A site blocks further tRNA entry, arresting elongation until the ribosome/mRNA complex is localized to the endoplasmic reticulum. The lesson from SRP function is that A-site occlusion is a viable and potent means to arrest translational elongation. Thus, any factor (protein or RNA) that can interact in or near the A site has the potential to stall elongation by blocking tRNA entry.

In this regard, there are many known factors that do interact with the A site—all for distinct reasons. The release factors, eRF1 and eRF3, the elongation factor EF-G, the ribosome recycling factors/mRNA decay factors DOM34 and HBS1, and the mRNA decay factor SKI7 are all thought to interact at the A site. Thus, the A site is a busy place and a potential target for mRNA-specific regulation. A-site occlusion could easily be achieved by a message-specific regulator provided it has sequence-specific binding properties for its mRNA transcript and a motif capable of A-site docking.

The SRP example demonstrates, in clear molecular terms, how elongation can be regulated. Importantly, the literature

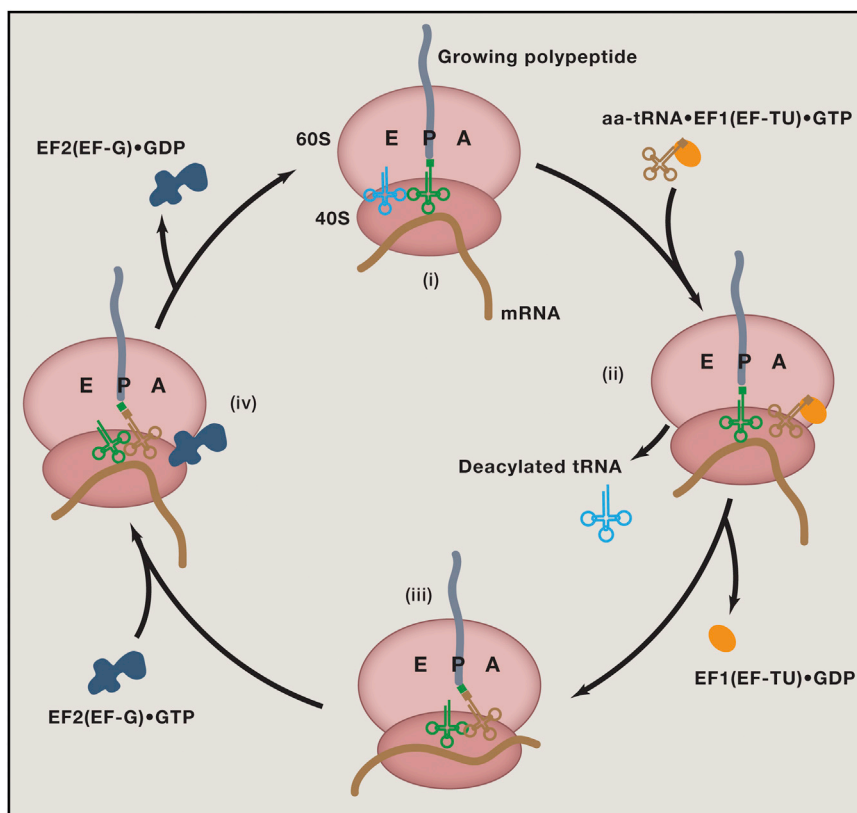


Figure 1. Translational Elongation at a Glance

Shown are the four basic steps of translational elongation. The ribosome has three major tRNA pockets, the A, P, and E sites. The first step of polypeptide elongation is the recognition and accommodation of the cognate tRNA, as directed by a mRNA codon, within the ribosomal A site (i). The cognate tRNA is brought into the ribosome as a complex with elongation factor 1 (EF-TU in bacteria) and GTP. Recognition of the cognate tRNA catalyzes the hydrolysis of GTP and the eviction of EF1 from the A site (ii). At this point, the deacylated tRNA in the E site is also thought to be evicted. The A-site tRNA and the P-site tRNA move into close proximity for the peptidyl transfer reaction, where the growing polypeptide chain is added to the amino acid on the A-site tRNA (iii). Elongation factor 2 (EF-G in bacteria) then enters the A site and completes ribosome translocation by moving the A-site tRNA to the P site and the P-site deacylated tRNA into the E site (iv). The process then repeats itself over and over again.

contains a number of less clear but still tantalizing glimpses of where modulating ribosome translocation might be a driving force in regulation. In several of these cases, regulated elongation is inferred from observations that repressed mRNA co-sediments with polysomes in sucrose gradients and/or that this co-sedimentation is resistant to puromycin treatment, which causes release of translating ribosomes. Nevertheless, accumulating evidence hints to a broad influence of translational elongation on the control of gene expression.

Control of Elongation during Early Development

Mechanisms of translational control during the early development of model organisms are often recapitulated in adult mammalian tissues. Consider, for example, the case of masked (i.e., repressed) mRNA in the oocytes (eggs) of sea urchins and frogs. As the oocytes prepare for fertilization during early embryogenesis, this mRNA is massively mobilized onto polysomes, which coincides with a substantial decrease in ribosome translocation time (Brandis and Raff 1978; Richter et al., 1982). These and other observations of this era now seem archaic because, for the most part, they pre-dated one's ability to assess the time required for a ribosome to transit any particular mRNA. Even so, they illustrate the point that polypeptide elongation rates can be regulated by cellular transitions.

Masked mRNA is also a hallmark of *Drosophila* development. Here, the translation of *nanos* mRNA, which encodes a posterior pole determinant, is regulated both spatially and temporally. Although *nanos* mRNA translation is controlled in multiple ways,

one is by ribosome stalling (Clark et al., 2000; Andrews et al., 2011). *Nanos* mRNA co-sediments with polysomes in sucrose gradients even though no *Nanos* protein is detected; yet, when the polysomes were added to an in vitro ribosome run-off system, the stalled polysomes resumed their transit and produced *Nanos* protein, showing that they were paused rather than immobilized in an inactive form. This scenario is somewhat similar to that observed with *oskar* mRNA, another posterior pole determinant in *Drosophila*. *Oskar* mRNA also co-sediments with polysomes even though little *Oskar* protein is detected (Braat et al., 2004). Moreover, when added to an in vitro translation system derived from ovaries, puromycin, an antibiotic that acts on translating ribosomes by mimicking tRNA and causing premature polypeptide release and ribosome dissociation, caused only a partial shift of the sedimentation of *oskar* mRNA to lighter fractions of sucrose gradients, suggesting that it is associated with stalled ribosomes.

Micro RNAs, which have profoundly changed our notion of how biological processes are regulated, were discovered during examination of *C. elegans* larval development (Lee et al., 1993; Wightman et al., 1993). In spite of the huge number of studies that have analyzed miRNA activity, the mechanism(s) by which they silence mRNA expression remains somewhat enigmatic, perhaps because they repress translation a number of different ways. That mRNAs silenced by miRNAs are often eventually destroyed is beyond doubt, but the step(s) at which the silencing occurs is seemingly manifold. Olsen and Ambros (1999) noted that although *C. elegans* *lin-4* miRNA inhibited *Lin-14* mRNA translation, the message appeared to be stable and co-sedimented with polysomes in sucrose gradients. This observation gave rise to the hypothesis that miRNAs repress translation by stalling ribosomes. Using cell lines, several labs subsequently found miRNA-inhibited mRNAs that also co-sedimented with polysomes (Nottrott et al., 2006; Maroney et al., 2006), and one

proposed that miRNAs promote pre-mature drop-off of translating ribosomes (Petersen et al., 2006). Although ribosome profiles derived from developing zebrafish embryos showed that at least miR-430 did not induce ribosome drop off (Bazzini et al., 2012), it remains an open question as to the extent to which miRNAs can promote post-initiation mRNA silencing.

Synaptic Plasticity in Neurons

Neuronal processes, particularly dendrites, have long been known to harbor mRNAs whose translation is critical for synaptic plasticity, the underlying cellular basis of learning and memory (Kang and Schuman 1996; Martin et al., 1997). The regulation of dendritic mRNA translation must be considered in conjunction with cellular localization as mRNAs are transported from cell bodies into dendrites on molecular motors in a mostly silent form. The mRNAs are then activated in response to synaptic activity (Kanai et al., 2004). Early work found that a substantial portion of neuronal mRNAs reside in granules that sediment in sucrose gradients to fractions much heavier than polysomes. Electron microscopy revealed these granules to be composed of densely packed polysomes, but because they lack the initiation factors eIF4E and eIF4G, they were thought to represent stalled ribosomes. Membrane depolarization of neurons by KCl-stimulated translation partially dispersed the aggregates, suggesting that the stalled polysomes resumed translation (Kriechvsky and Kosik 2001).

In a contemporaneous study, Scheetz et al. (2000) stimulated synaptoneurosomes (a biochemical preparation of pre- and post-synaptic compartments) isolated from rat brain with the neurotransmitter N-methyl-D-aspartate (NMDA) and observed an increase in eEF2 phosphorylation, which would inactivate the enzyme and thus slow ribosome translocation. These investigators found a concurrent increase in the synthesis of the critical synaptic protein alpha calcium/calmodulin protein kinase II (α CaMKII) and hypothesized that inhibition of elongation of some mRNAs allows for the elevated translation of other mRNAs by mechanisms involving enhanced initiation. This hypothesis suggests that phospho-eEF2 could discriminate among mRNAs, which might be accomplished by, for example, spatial segregation of some components of the translational apparatus.

A link between eEF2 phosphorylation and synaptic activity was also observed by Sutton et al. (2007), who found that eEF2's enzymatic activity can be toggled by the type of neurotransmission to which a neuron is subjected. eEF2 is mostly non-phosphorylated when neurotransmission is evoked by action potentials. On the other hand, miniature synaptic transmission, which is spontaneous in nature and independent of action potentials, results in eEF2 phosphorylation. As might be expected, protein synthesis is up or downregulated depending on the state of eEF2 phosphorylation. For this reason, Sutton et al. (2007) proposed that eEF2 is a sensor that links synaptic activity to local (i.e., dendritic) control of polypeptide elongation.

More contemporary studies are consistent with these findings. Using an indirect in vivo ribosome run-off assay in neurons, Graber et al. (2013) found that mRNAs in dendrites are associated with stalled polysomes that can be reactivated by stimulation of metabotropic glutamate receptors, which induces long-term depression (LTD), a protein synthesis-dependent form of

synaptic plasticity. Buxbaum et al. (2014) used three-color fluorescence in situ hybridization (FISH) and single-molecule detection technology to explore actin mRNA masking/unmasking dynamics in neuronal dendrites. Their results buttress the interpretation that quiescent mRNAs in neurons associate with stalled ribosomes and that induction of long-term potentiation (LTP), another form of synaptic plasticity, activates the stalled ribosomes to complete translation.

Ribosome Stalling in Fragile X Syndrome

Fragile X syndrome is the most common form of inherited intellectual disability and most frequent monogenic cause of autism. The syndrome is caused by a CGG repeat expansion in the *Fmr1* gene, which causes its transcriptional inactivation. *Fmr1* encodes FMRP, an RNA binding protein that represses translation. In the absence of FMRP, protein synthesis in the brain is excessive, and it is commonly thought that this leads to synaptic dysfunction and other anomalies associated with this disease.

Several studies have shown that FMRP co-sediments with polysomes, suggesting that it inhibits translation by stalling ribosomes (Feng et al., 1997; Corbin et al., 1997; Stefani et al., 2004). In a groundbreaking study by Darnell et al. (2011), FMRP was found to crosslink to nearly 1,000 mRNAs in the brain in an experiment where UV irradiation to covalently link RNAs and proteins is followed by immunoprecipitation and deep sequencing (this procedure is referred to as "CLIP"). Surprisingly, most of the sites in mRNA to which FMRP was crosslinked were distributed in coding regions in a *cis* element-independent manner. Moreover, extensive analysis of the associated mRNAs showed them to co-sediment with polysomes but to be largely resistant to puromycin treatment (because puromycin causes dissociation of transiting, but not static ribosomes, one infers that ribosomes do not move, or move only very slowly, on FMRP-bound mRNAs). By contrast, mRNAs not crosslinked to FMRP dissociated from polysomes following puromycin treatment and sedimented with the nontranslating ribonucleoprotein (RNP) fractions of sucrose gradients. Therefore, a synthesis of these two observations—that FMRP binds to coding regions and that the ribosomes on such mRNAs do not transit—strongly implies that FMRP stalls ribosomes.

In a number of cases, mouse models have shown that the excessive protein synthesis in the Fragile X brain can be restored to normal levels if a second gene is deleted, which leads to rescue of many disease phenotypes (Dölen et al., 2007; Bhattacharya et al., 2012; Udagawa et al., 2013; Gross et al., 2015). For example, Udagawa et al. (2013) focused on CPEB1, which generally (although not necessarily exclusively) stimulates translation in the brain and co-localizes and co-immunoprecipitates with FMRP. The binding site for CPEB1 is present in about 30% of the FMRP-associated mRNAs and thus has the potential to regulate their expression. These authors proposed that translational homeostasis in the brain might be restored if the *CPEB1* gene, as well as *Fmr1*, was disrupted. Indeed, not only was the excessive protein synthesis restored to normal levels in FMRP/*CPEB* double-knockout mice, but so too were a variety of pathologies associated with Fragile X. Moreover, Udagawa et al. (2013) found that ribosome transit time in the Fragile X brain

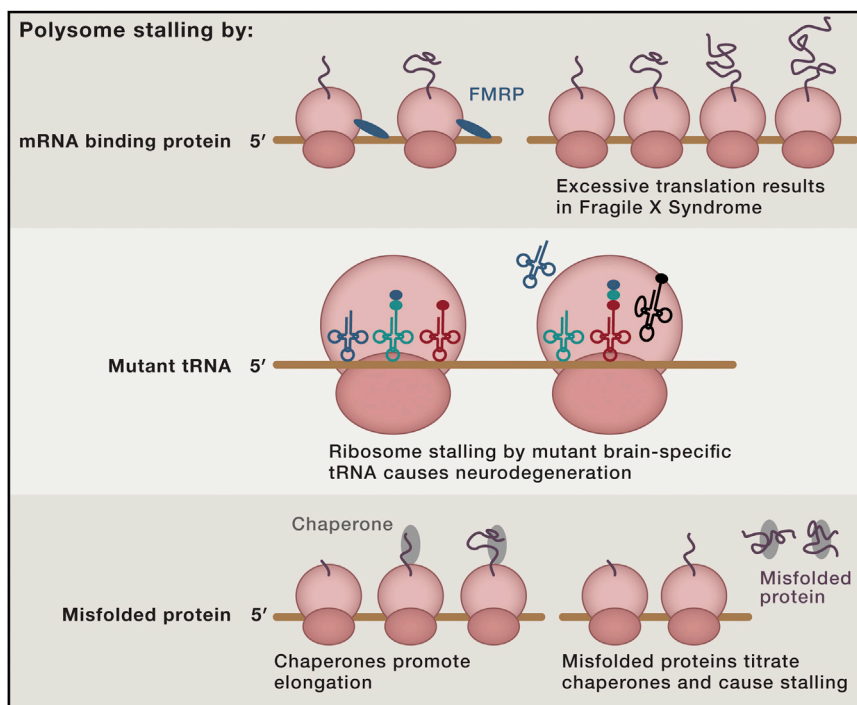


Figure 2. Three Examples of Regulated Polypeptide Elongation

FMRP is proposed to bind both the ribosome and the engaged mRNA to impede ribosome translocation. FMRP is not produced when the *Fmr1* gene is inactivated, which results in an elevated rate of polypeptide elongation and the Fragile X syndrome (top). In the brain, a mutated tissue-specific tRNA causes ribosome stalling at its corresponding codon. Neurodegeneration occurs if GTPBP2 is also mutated (middle). During proteotoxic stress, the chaperone HSC70, which normally binds the nascent peptide as it emerges from the ribosome, is titrated by misfolded proteins and causes ribosome stalling after reading about 50 codons (bottom).

control is becoming a hallmark of neural development and function. One stunning observation on translational control at the level of ribosome translocation came from a mutagenesis screen for neurological disorders in mice where one line in particular displayed profound brain degeneration, ataxia, and death by about 2 months of age (Ishimura et al., 2014). Mapping the mutation by

was ~45% faster than in the wild-type brain, which was rescued to normal in the FMRP/CPEB1 double-knockout brain. How the loss of CPEB1 slows ribosome translocation in the FMRP-deficient brain is unknown, but irrespective of the mechanism involved, these data directly demonstrate that FMRP controls polypeptide elongation and that Fragile X may be a disease, at least in part, of accelerated ribosome translocation.

Because the ribosome has intrinsic helicase-like activity and is capable of removing protein/RNA complexes as it moves along the mRNA (Takyar et al., 2005), FMRP acting as a simple roadblock to stall ribosomes, while tenable as a model, seems too simplistic. An intriguing structural study of *Drosophila* FMRP showed that it binds directly to the ribosome via an interaction with ribosomal protein L5 (Chen et al., 2014). The two KH (hnRNP K homology) domains of FMRP interact with the ribosome, while the single RGG (arginine-glycine-glycine) box could associate directly with RNA, suggesting that FMRP may act as a bridge between the two to impede ribosome movement (Figure 2). However, such a configuration does not predict whether the specificity of FMRP binding is imparted by its association with the mRNA or the ribosome.

One extant question arising from the above discussion is whether FMRP-mediated translational repression is permanent or reversible. At least one form of synaptic plasticity, LTD, induces FMRP phosphorylation and ubiquitin-mediated destruction (Nalavadi et al., 2012; Huang et al., 2015), which one surmises would remove the block to ribosome translocation and allow polypeptide elongation to proceed.

Ribosome Stalling in Neurodegeneration

The brain contains the most complex mixture of mRNAs of any adult tissue, and therefore, it is not surprising that translational

crosses to congenic mouse strains combined with SNP analysis showed that the mutagen (N-ethyl-N-nitrosourea) produced a point mutation in a splice site of the *Gtpbp2* gene, which encodes a protein with homology to GTPases involved in translation, particularly the ribosome recycling factors HBS1 and eRF3. These two proteins interact with the ribosome release factors Dom34 (yeast nomenclature) and eRF1, and GTPBP2 does indeed co-immunoprecipitate with Pelota, the mammalian homolog of Dom34. However, neurodegeneration was manifest only in the commonly used C57BL/6J background, suggesting that a second modifier gene specific to this strain was necessary to produce the phenotype. The modifier gene was found to encode an arginine tRNA isodecoder (isodecoder tRNAs share the same anticodon but have changes elsewhere in the molecule). This arginine tRNA_{UCU}, which contains a single C to U mutation in the T-stem loop that is likely to result in RNA misfolding, is CNS specific; in contrast, GTPBP2 is widely expressed. Ishimura et al. (2014) hypothesized that the mutant tRNA^{Arg}_{UCU} would cause the ribosome to stall at AGA codons. Indeed, ribosome profiling (Ingolia et al., 2009) revealed a particularly high number of reads at AGA codons in cerebellar material containing the mutant tRNA, indicating strong ribosome stalling at these sites. Such strong stalling was not evident in cerebellar tissue from animals with wild-type tRNA^{Arg}_{UCU} (Ishimura et al., 2014). Thus, ribosome stalling at AGA codons caused by the mutant tRNA may allow mutant GTPBP2 to recruit Pelota and promote premature polypeptide release, resulting in neurodegeneration (Figure 2) (Ishimura et al., 2014; Darnell 2014). tRNA mis-charging (Lee et al., 2006) or mis-folding appears to have particularly profound consequences for CNS function, suggesting that disruption of tRNA activity could contribute to other CNS pathologies as well.

Polysome Pausing by Misfolded Proteins

Cellular stress has long been known to cause reduced translation and promote the formation of stress granules and processing bodies (p bodies) where mRNAs have been hypothesized to undergo silencing and decay, respectively (Anderson and Kedersha 2009; Decker and Parker 2012). Proteotoxicity can be one source of cellular stress and can be simulated by feeding cells with amino acid analogs that result in misfolded proteins. Liu et al. (2013) found that culturing cells in L-azetidine-2-carboxylic acid (AZC), a proline analog, caused rapid turnover of newly synthesized protein, and when AZC treatment was combined with the proteasome inhibitor MG132, protein synthesis substantially decreased. This treatment did not induce rapid stress granule formation or eIF2 α phosphorylation, which are common stress-induced events (Proud, 2005), nor did it alter other parameters normally associated with a block at initiation. Instead, AZC and MG132 caused polypeptide elongation to slow. Ribosome profiling showed that the stress promoted ribosome stalling within ~50 codons of the initiation AUG, enough to encode a polypeptide partially buried in the exit tunnel of the ribosome and partially exposed to the cellular milieu. Liu et al. (2013) hypothesized that, because the chaperone HSC70 binds and helps fold nascent peptides as they emerge from the ribosome, it would naturally facilitate ribosome transit. During proteotoxic stress, however, the chaperone may be titrated by misfolded protein and thus cause ribosome stalling and reduced translation (Figure 2).

Heat shock, perhaps of the most common form of stress, has been known for decades to inhibit protein synthesis at least partly at the level of elongation (Ballinger and Pardue 1983). Using ribosome profiling, Shalgi et al. (2013) found that heat stress induces widespread stalling at about codon 65. Similar to the data presented by Liu et al. (2013), this study found that HSP70 was responsible for the ribosomal stall. Although Shalgi et al. (2013) did not determine how the chaperone induces stalling during heat stress, they suggest that it may involve HSP70 association with the emergent nascent peptide, that it somehow clogs the peptide exit tunnel, or that it mitigates elongation factor activity.

An additional mode by which heat shock proteins modulate elongation comes out of structural analysis of yeast ribosomes (Zhang et al., 2014). A ribosome-associated complex (RAC), which is composed of HSP40 and HSP70, binds both ribosomal subunits through a single long α helix. As a consequence, a necessary rotation between the subunits during translation is limited and thus ribosome translocation is reduced. Zhang et al. (2014) propose that RAC somehow responds to the folding needs of the nascent peptide as it emerges from the exit tunnel and binds the ribosome to reduce translocation such that the peptide can assume a productive tertiary structure.

Caloric restriction or nutrient deprivation also induces cell stress. Although a number of proteins can sense that nutrients are in short supply, the major one is mTORC1 (mTOR complex 1). When amino acid levels are low, this kinase is inactivated, which leads to a number of downstream dephosphorylation events, including that of 4EBP1. The non-phosphorylated form of this protein disrupts the ability of eIF4E to bind eIF4G and recruit factors necessary to initiate translation. A second nutrient

sensor is AMP kinase (AMPK). In response to amino acid starvation, it downregulates protein synthesis and other high ATP-demanding processes in an effort to conserve energy. One way AMPK accomplishes this task is by activating eEF2 kinase (eEF2K), which reduces elongation by phosphorylating eEF2 (Leprivier et al., 2013). Through this kind of energy conservation, AMPK-mediated eEF2 phosphorylation promotes cell survival in the face of nutrient restriction.

Perhaps the most unusual but well-defined example of stress-regulated elongation is yeast *Hac1* mRNA, which encodes a transcription factor involved in the unfolded protein response (UPR). As the 5' end of *Hac1* mRNA emerges, it associates with ribosomes and begins to be translated. When the 3' UTR enters the cytoplasm, an intron contained within it base pairs with the 5'UTR, thereby forming a closed loop that stalls the ribosomes. In response to ER stress, the 3' UTR intron is removed by the nuclease Ire1p and the RNA ends are unconventionally spliced by tRNA ligase, which consequently allows the ribosomes to continue to catalyze polypeptide elongation (Chapman and Walter 1997; Rügsegger et al., 2001).

Regulated Elongation during Oncogenic Transformation

The AMPK-eEF2 kinase pathway mediates tumorigenesis, as well as caloric-restriction-induced stress. When tumor cells with low levels of eEF2K are starved, polypeptide elongation remains robust and energy is consumed at a high rate, thereby causing the cells to undergo apoptosis. However, some tumor cells can adapt to starvation conditions by upregulating the AMPK-eEF2 pathway, which inhibits polypeptide elongation, conserves energy, and promotes cell survival.

mTOR is a master regulator of cell physiology and can induce many types of cancer through phosphorylation of 4EBP1 and other substrates. As noted previously, 4EBP1 inhibits initiation by competitively binding eIF4E to the exclusion of eIF4G. When 4EBP1 is phosphorylated, it dissociates from eIF4E, which allows eIF4E, eIF4G, and eIF4A to form the eIF4F initiation complex. As a consequence of these events, mRNAs with particularly long and complex 5' UTRs are preferentially translated because eIF4A is an RNA helicase that unwinds RNA secondary structure. Many mRNAs encoding oncogenes or growth-promoting factors have complex 5' UTRs and are upregulated by mTOR and promote cellular transformation (Pelletier et al., 2015). mTOR can also modulate polypeptide elongation to facilitate cell proliferation and cancer etiology. In intestinal cancer caused by mutations in adenomatous polyposis coli (APC), a tumor suppressor, mTORC1 can modulate translation via phosphorylation of 4EBP1 or S6 kinase (S6K), the latter of which is an upstream inactivator of eEF2K. Thus, at least for this cancer type, tumor growth is regulated by an axis of mTORC1-S6K-eEF2K-eEF2, which culminates in decreased elongation rates that are sensitive to the levels of S6K, but not 4EBP1 (Faller et al., 2015).

Regulation by tRNAs

In the previous examples, we have seen that mRNA-specific regulation might be achieved by turning translational elongation on or off. In addition to these specialized cases, some evidence hints that the control of translation elongation might broadly occur and be critical for gene regulation. Unlike the

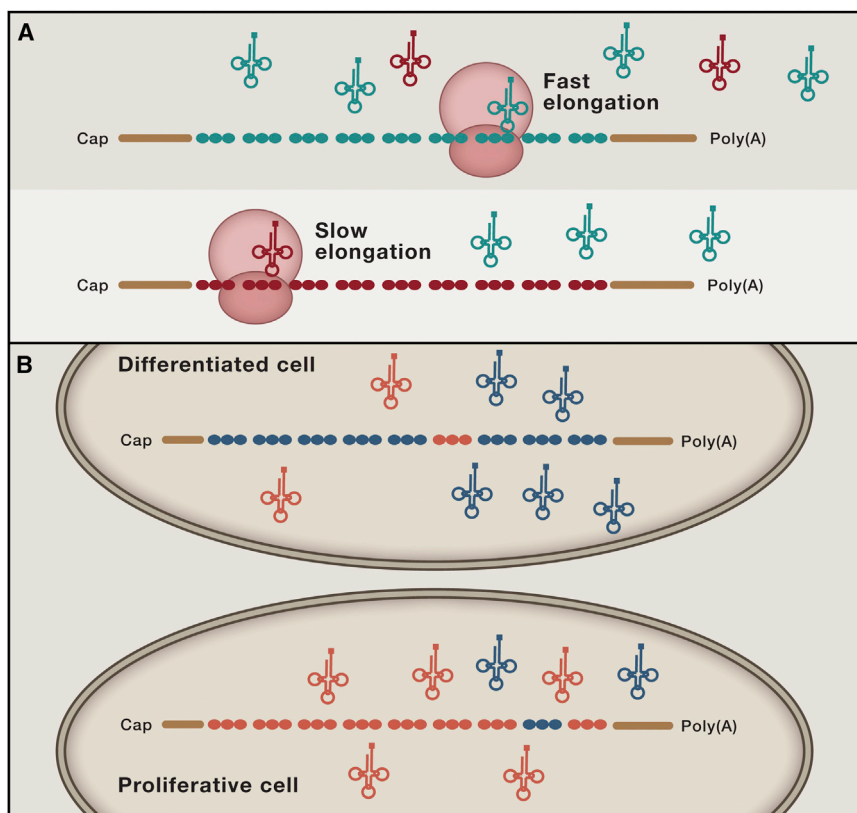


Figure 3. Illustration of the Concept of Optimality

(A) Illustration of codon optimality showing two hypothetical mRNAs. The green circles represent optimal codons, whereas the red circles represent non-optimal codons. Designation of codons as optimal or non-optimal is a function of the concentration of tRNA in the cell. The green tRNA concentrations are high, whereas the red tRNA concentrations are low; this difference impacts the speed of elongation.

(B) Differentiated and proliferative cells have varied concentrations of tRNAs that are tailored to the decoding requirements of the expressed mRNAs. In this example, differentiated cells have an abundance of certain types of tRNAs (depicted in blue) because the mRNAs they express require similarly high levels of the corresponding codons. Conversely, proliferating cells contain different sets of tRNAs (in orange) to match the codons in mRNAs enriched in these cells (see [Gingold et al., 2014](#)).

aforementioned binary switches, ribosome translocation rates may subtly influence the expression of all messages. If translocation rates are mRNA specific, then elongation would play a pivotal role in the synthesis, folding, and perhaps function of all proteins. The hypothesis that each mRNA has a distinct elongation rate is based on the notion of supply and demand: supply of functional tRNAs and demand by expressed codons. In this light, tRNAs are implicated as critical regulators of the expressed transcriptome.

In 1968, Francis Crick referred to the degeneracy of the genetic code as a “frozen accident” based on the required 64 combinations needed to code for 20 amino acids ([Crick, 1968](#)). Since then, a prevailing zeitgeist has been that synonymous codon substitutions are silent, having no bearing on gene function. Antithetically, a growing body of literature suggests that synonymous codons are differentially recognized by the translational apparatus. This concept has been referred to as codon optimality ([Reis et al., 2004](#); [Novoa and Ribas de Pouplana, 2012](#); [Pechmann and Frydman, 2013](#); [Krisiko et al., 2014](#)). Codon optimality should not be confused with codon usage or bias. Codon usage/bias is the overrepresentation of certain codons within the genome and is the result of numerous selective pressures, including translational elongation rate, translational accuracy, splicing, and 5' UTR structure ([Akashi, 1994](#); [Parmley et al., 2006](#); [Drummond and Wilke, 2008](#); [Gu et al., 2010](#)). The term codon optimality has been introduced in an attempt to specifically define the differential recognition of codons by the translational apparatus and is a property that is distinct from the usage

of codons within the genome. For instance, commonly occurring codons can be classified as optimal or non-optimal, whereas uncommon codons can also be optimal or non-optimal with respect to their influence on translational elongation rate ([Presnyak et al., 2015](#)).

Conceptually, codon optimality reflects the balance between the supply of charged tRNA molecules and their demand

imposed by the concentration of codons engaged in translation ([Figure 3](#)). Thus, the ribosome decodes some codons quickly because their cognate tRNAs are abundant, whereas other codons are read more slowly because their tRNA concentrations are more limiting ([Tuller et al., 2010](#); [Novoa and Ribas de Pouplana, 2012](#)). In addition, codon optimality is based somewhat on the accuracy of tRNA anticodon/codon interactions, which can influence decoding rate ([Akashi, 1994](#); [Drummond and Wilke, 2008](#)). The theory that each codon is read by the ribosome at subtly distinct rates would predict that the kinetics of protein synthesis are determined by the primary sequence of every gene ([Presnyak et al., 2015](#)). Thus, the overall elongation rate is the sum of each codon's infinitesimally small effect on ribosome translocation. Recent support for this hypothesis has come from several labs demonstrating that codon optimality is a powerful determinant of both mRNA translation elongation and mRNA stability, which are tightly coupled events ([Pechmann and Frydman, 2013](#); [Presnyak et al., 2015](#)).

If codon content dictates translational elongation rate, then perhaps it can be regulated by changing the functional concentration of tRNAs. If tRNA pools change in response to stress, environmental conditions, or other biological cues, then so would the rate at which a codon is read and thus ultimately ribosome translocation rates. Recent studies demonstrate that tRNA levels and modifications indeed fluctuate in response to biological cues. First, [Gingold et al. \(2014\)](#) demonstrated that tRNA pools fluctuate in over 470 tumor samples when compared to quiescent cells. Specifically, a subset of tRNAs is induced in

proliferating cells and repressed in quiescent cells. Moreover, a distinct subset of tRNAs is repressed in proliferating cells that are active in quiescent cells. Importantly, the tRNAs that are induced during proliferation often have anticodons corresponding to codons enriched in cell-autonomous genes. In contrast, tRNAs induced in differentiated cells carry anticodons for codons enriched in mRNAs for cell adhesion, cell-junction assembly, toll-like receptor signaling, and extracellular matrix genes. In a study in which mouse embryonic development was examined, [Schmitt et al. \(2014\)](#) demonstrated that tRNA expression patterns were controlled to generate an anticodon pool that corresponds to the codon demand by mRNAs ([Schmitt et al., 2014](#)). Thus, the patterns of tRNA expression match codon usage within the expressed transcriptome ([Figure 3](#)).

Stress can also alter the level of functional tRNAs within a cell. Specifically, reprogramming of tRNA modifications occurs in cells exposed to different conditions. [Chan et al. \(2012\)](#) demonstrated that exposure of cells to hydrogen peroxide results in an increase in the amount of tRNA^{Leu(CAA)} containing 5-methylcytosine (m⁵C) at the wobble position. This increase in m⁵C causes selective translation of mRNAs enriched in the TTG codon. A nutrient-driven tRNA modification has also been observed in *Drosophila*. Here, the bioavailability of queuine (a modified base) affects the levels of queuosine-modified tRNAs ([Zaborske et al., 2014](#)). Queuine is scavenged by eukaryotes from the tRNAs of bacteria and absorbed in the gut where, at least in flies, it alters translation profiles. Together, these data suggest that tRNA pools can and do fluctuate in response to biological cues. If these concepts occur more broadly, then it places tRNAs as important and under-appreciated regulators of mRNA post-transcriptional regulation.

Why Regulate at Elongation?

The regulation of translational elongation might afford several advantages to an mRNA and the cell. First, loading an mRNA onto polyribosomes and then stalling the polysomes in response to the presence or absence of a stimulus would allow instantaneous production of new polypeptides once the stimulus was changed. This rapid response might be especially important in situations where, as in neurons, immediate protein synthesis is needed in response to synaptic stimulation. Lodging mRNA on translationally quiescent polyribosomes might also serve a protective function, limiting access by nucleases. In the cell, RNAs tend to be degraded when not associated with protein factors; therefore, unless there are active events reorganizing the mRNA out of translation and into a translation-repressed ribonucleoprotein complex (e.g., stress granules or maternal mRNA storage granules), polyribosomes might serve a protective role.

Basal regulation of elongation rate by tRNAs might in fact be a primary driver of protein levels within the cell. It was recently observed that tightly coordinated optimal codon content occurs in genes encoding proteins with common physiological function ([Presnyak et al., 2015](#)). This finding suggests that there is evolutionary pressure toward certain synonymous codon usage to coordinate gene expression at the level of protein synthesis and mRNA decay. The coordination of protein complexes through coordinate codon-dependent elongation rates would provide

an elegant mechanism to ensure a consistent stoichiometric relationship between all members of a given complex. Because this coordination is based on codon choice, changing tRNA levels and/or modifications would provide a simple yet sophisticated means to uniformly regulate an entire physiological process by changing ribosome elongation rates.

The Road Ahead

The advent of ribosome profiling ([Ingolia et al., 2009](#)), which displays stalled ribosomes on an mRNA-specific and codon-specific basis, has firmly placed elongation on the map of important gene regulatory mechanisms. Although, in most cases, the salient molecular details for stalling are just beginning to come into focus, perhaps some general principles may be involved. One is exemplified by FMRP in which a protein acts on a group of mRNAs to stall ribosome translocation. Certainly, intrinsic mRNA sequence gives specificity to the stalling, but is there also something intrinsic to the ribosome that allows it to be stalled? For example, we know the ribosome has trouble with lysine AAA codons ([Koutmou et al., 2015](#)) and that polyproline in the exit tunnel can also slow elongation ([Gutierrez et al., 2013](#)). Thus, certain ribosome characteristics could be exploited by an mRNA to slow elongation, and consolidation of this event could occur through a second factor such as FMRP. Extrapolation of the data from several studies on FMRP suggests that this could be the case and that both message and ribosome contribute to the stalling.

A second general principle centers on tRNA. The astonishing study of [Ishimura et al. \(2014\)](#) shows that the brain contains a tRNA that is absent from other tissues. When this tRNA has a single base change, ribosomes stall at its cognate mRNA codon. This stall results in CNS-specific pathology in instances where GTPBP2 is also mutated. Is tRNA specificity widespread, and if so, how does it influence health and disease? A corollary of tRNA tissue specificity is “optimality”—the matchup between codon prevalence and relative of abundance of the tRNA bearing the anticodon. If this ratio becomes skewed during development or times of stress, it is easy to see how it could alter ribosome translocation and result in a biological transition.

Therapeutics aimed at the translational landscape have, for the most part, concentrated on initiation (e.g., [Pelletier et al., 2015](#)). Importantly, however, just this year, a novel multiple-stage antimalarial agent was discovered whose target is Elongation Factor 2 ([Baragaña et al., 2015](#)). Thus, we wonder whether the elongation phase of protein synthesis is an equally promising target for therapeutic intervention to ameliorate disease.

ACKNOWLEDGMENTS

We thank Drs. Lori Lorenz, Botao Liu, and Sophie Martin for comments on the manuscript. Work from the authors' laboratories was supported by grants from the National Institutes of Health (R01 GM46779, R01 NS079415, and U54 082013 to J.D.R. and R01 GM080465 to J.C.).

REFERENCES

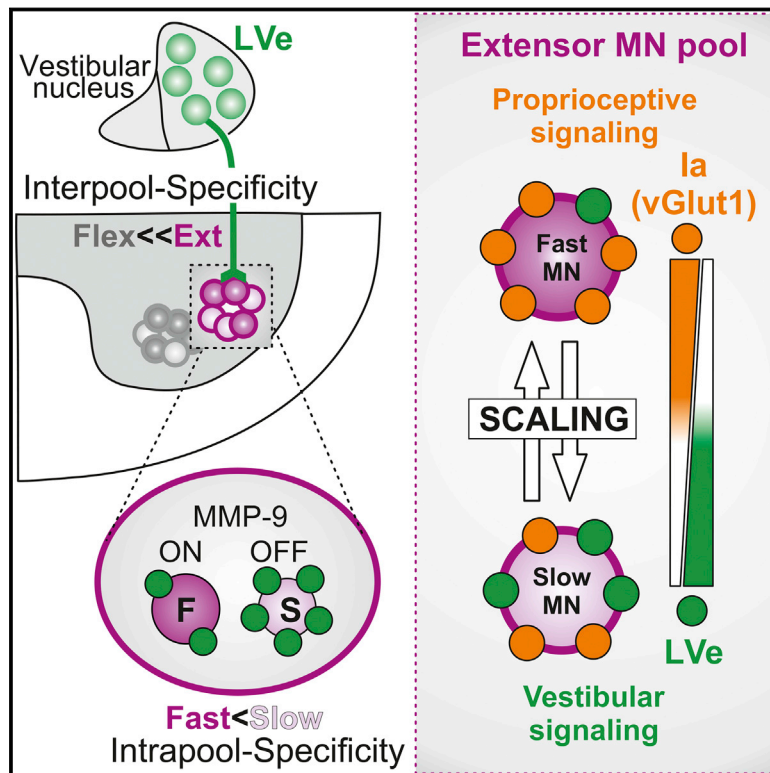
Akashi, H. (1994). Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* 136, 927–935.

- Anderson, P., and Kedersha, N. (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat. Rev. Mol. Cell Biol.* 10, 430–436.
- Andrews, S., Snowflack, D.R., Clark, I.E., and Gavis, E.R. (2011). Multiple mechanisms collaborate to repress nanos translation in the *Drosophila* ovary and embryo. *RNA* 17, 967–977.
- Ballinger, D.G., and Pardue, M.L. (1983). The control of protein synthesis during heat shock in *Drosophila* cells involves altered polypeptide elongation rates. *Cell* 33, 103–113.
- Baragaña, B., Hallyburton, I., Lee, M.C.S., Norcross, N.R., Grimaldi, R., Otto, T.D., Proto, W.R., Blagborough, A.M., Meister, S., Wirjanata, G., et al. (2015). A novel multiple-stage antimalarial agent that inhibits protein synthesis. *Nature* 522, 315–320.
- Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* 336, 233–237.
- Bhattacharya, A., Kaphzan, H., Alvarez-Dieppa, A.C., Murphy, J.P., Pierre, P., and Klann, E. (2012). Genetic removal of p70 S6 kinase 1 corrects molecular, synaptic, and behavioral phenotypes in fragile X syndrome mice. *Neuron* 76, 325–337.
- Braat, A.K., Yan, N., Arn, E., Harrison, D., and Macdonald, P.M. (2004). Localization-dependent oskar protein accumulation; control after the initiation of translation. *Dev. Cell* 7, 125–131.
- Brandis, J.W., and Raff, R.A. (1978). Translation of oogenetic mRNA in sea urchin eggs and early embryos. Demonstration of a change in translational efficiency following fertilization. *Dev. Biol.* 67, 99–113.
- Buxbaum, A.R., Wu, B., and Singer, R.H. (2014). Single β -actin mRNA detection in neurons reveals a mechanism for regulating its translatability. *Science* 343, 419–422.
- Chan, C.T.Y., Pang, Y.L.J., Deng, W., Babu, I.R., Dyavaiah, M., Begley, T.J., and Dedon, P.C. (2012). Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nat. Commun.* 3, 937.
- Chapman, R.E., and Walter, P. (1997). Translational attenuation mediated by an mRNA intron. *Curr. Biol.* 7, 850–859.
- Chen, E., Sharma, M.R., Shi, X., Agrawal, R.K., and Joseph, S. (2014). Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Mol. Cell* 54, 407–417.
- Clark, I.E., Wyckoff, D., and Gavis, E.R. (2000). Synthesis of the posterior determinant Nanos is spatially restricted by a novel cotranslational regulatory mechanism. *Curr. Biol.* 10, 1311–1314.
- Corbin, F., Bouillon, M., Fortin, A., Morin, S., Rousseau, F., and Khandjian, E.W. (1997). The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. *Hum. Mol. Genet.* 6, 1465–1472.
- Crick, F.H. (1968). The origin of the genetic code. *J. Mol. Biol.* 38, 367–379.
- Darnell, J.C. (2014). Molecular biology. Ribosome rescue and neurodegeneration. *Science* 345, 378–379.
- Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–261.
- Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb. Perspect. Biol.* 4, a012286.
- Dölen, G., Osterweil, E., Rao, B.S., Smith, G.B., Auerbach, B.D., Chattarji, S., and Bear, M.F. (2007). Correction of fragile X syndrome in mice. *Neuron* 56, 955–962.
- Drummond, D.A., and Wilke, C.O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134, 341–352.
- Faller, W.J., Jackson, T.J., Knight, J.R., Ridgway, R.A., Jamieson, T., Karim, S.A., Jones, C., Radulescu, S., Huels, D.J., Myant, K.B., et al. (2015). mTORC1-mediated translational elongation limits intestinal tumour initiation and growth. *Nature* 517, 497–500.
- Feng, Y., Absher, D., Eberhart, D.E., Brown, V., Malter, H.E., and Warren, S.T. (1997). FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. *Mol. Cell* 1, 109–118.
- Gingold, H., Tehler, D., Christoffersen, N.R., Nielsen, M.M., Asmar, F., Kooistra, S.M., Christophersen, N.S., Christensen, L.L., Borre, M., Sorensen, K.D., et al. (2014). A dual program for translation regulation in cellular proliferation and differentiation. *Cell* 158, 1281–1292.
- Graber, T.E., Hébert-Seropian, S., Khoutorsky, A., David, A., Yewdell, J.W., Lacaille, J.C., and Sossin, W.S. (2013). Reactivation of stalled polyribosomes in synaptic plasticity. *Proc. Natl. Acad. Sci. USA* 110, 16205–16210.
- Gross, C., Raj, N., Molinaro, G., Allen, A.G., Whyte, A.J., Gibson, J.R., Huber, K.M., Gourley, S.L., and Bassell, G.J. (2015). Selective role of the catalytic PI3K subunit p110 β in impaired higher order cognition in fragile X syndrome. *Cell Rep.* 11, 681–688.
- Gu, W., Zhou, T., and Wilke, C.O. (2010). A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS Comput. Biol.* 6, e1000664.
- Gutierrez, E., Shin, B.-S., Woolstenhulme, C.J., Kim, J.-R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. *Mol. Cell* 51, 35–45.
- Halic, M., Becker, T., Pool, M.R., Spahn, C.M.T., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427, 808–814.
- Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.* 83, 779–812.
- Huang, J., Ikeuchi, Y., Malumbres, M., and Bonni, A. (2015). A Cdh1-APC/FMRP ubiquitin signaling link drives mGluR-dependent synaptic plasticity in the mammalian brain. *Neuron* 86, 726–739.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223.
- Ishimura, R., Nagy, G., Dotu, I., Zhou, H., Yang, X.L., Schimmel, P., Senju, S., Nishimura, Y., Chuang, J.H., and Ackerman, S.L. (2014). RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science* 345, 455–459.
- Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* 43, 513–525.
- Kang, H., and Schuman, E.M. (1996). A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* 273, 1402–1406.
- Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., and Green, R. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. *eLife* 4. <http://dx.doi.org/10.7554/eLife.05534>.
- Krichevsky, A.M., and Kosik, K.S. (2001). Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation. *Neuron* 32, 683–696.
- Krisko, A., Copic, T., Gabaldón, T., Lehner, B., and Supek, F. (2014). Inferring gene function from evolutionary change in signatures of translation efficiency. *Genome Biol.* 15, R44.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854.
- Lee, J.W., Beebe, K., Nangle, L.A., Jang, J., Longo-Guess, C.M., Cook, S.A., Davisson, M.T., Sundberg, J.P., Schimmel, P., and Ackerman, S.L. (2006). Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature* 443, 50–55.
- Leprieux, G., Remke, M., Rotblat, B., Dubuc, A., Mateo, A.R., Kool, M., Agnihotri, S., El-Naggar, A., Yu, B., Somasekharan, S.P., et al. (2013). The eEF2 kinase confers resistance to nutrient deprivation by blocking translation elongation. *Cell* 153, 1064–1079.

- Liu, B., Han, Y., and Qian, S.B. (2013). Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. *Mol. Cell* 49, 453–463.
- Mamane, Y., Petroulakis, E., Rong, L., Yoshida, K., Ler, L.W., and Sonenberg, N. (2004). eIF4E—from translation to transformation. *Oncogene* 23, 3172–3179.
- Maroney, P.A., Yu, Y., Fisher, J., and Nilsen, T.W. (2006). Evidence that micro-RNAs are associated with translating messenger RNAs in human cells. *Nat. Struct. Mol. Biol.* 13, 1102–1107.
- Martin, K.C., Casadio, A., Zhu, H., Yaping, E., Rose, J.C., Chen, M., Bailey, C.H., and Kandel, E.R. (1997). Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell* 91, 927–938.
- Nalavadi, V.C., Muddashetty, R.S., Gross, C., and Bassell, G.J. (2012). Dephosphorylation-induced ubiquitination and degradation of FMRP in dendrites: a role in immediate early mGluR-stimulated translation. *J. Neurosci.* 32, 2582–2587.
- Nottrott, S., Simard, M.J., and Richter, J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat. Struct. Mol. Biol.* 13, 1108–1114.
- Novoa, E.M., and Ribas de Pouplana, L. (2012). Speeding with control: codon usage, tRNAs, and ribosomes. *Trends Genet.* 28, 574–581.
- Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* 216, 671–680.
- Parmley, J.L., Chamary, J.V., and Hurst, L.D. (2006). Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. *Mol. Biol. Evol.* 23, 301–309.
- Pechmann, S., and Frydman, J. (2013). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* 20, 237–243.
- Pelletier, J., Graff, J., Ruggero, D., and Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. *Cancer Res.* 75, 250–263.
- Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* 21, 533–542.
- Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., and Collier, J. (2015). Codon optimality is a major determinant of mRNA stability. *Cell* 160, 1111–1124.
- Proud, C.G. (2005). eIF2 and the control of cell physiology. *Semin. Cell Dev. Biol.* 16, 3–12.
- Reis, M.D., Savva, R., and Wernisch, L. (2004). Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32, 5036–5044.
- Richter, J.D., and Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477–480.
- Richter, J.D., Wasserman, W.J., and Smith, L.D. (1982). The mechanism for increased protein synthesis during *Xenopus* oocyte maturation. *Dev. Biol.* 89, 159–167.
- Rüegsegger, U., Leber, J.H., and Walter, P. (2001). Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell* 107, 103–114.
- Scheetz, A.J., Nairn, A.C., and Constantine-Paton, M. (2000). NMDA receptor-mediated control of protein synthesis at developing synapses. *Nat. Neurosci.* 3, 211–216.
- Schmitt, B.M., Rudolph, K.L.M., Karagianni, P., Fonseca, N.A., White, R.J., Talianidis, I., Odom, D.T., Marioni, J.C., and Kutter, C. (2014). High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res.* 24, 1797–1807.
- Shalgi, R., Hurt, J.A., Krykbaeva, I., Taipale, M., Lindquist, S., and Burge, C.B. (2013). Widespread regulation of translation by elongation pausing in heat shock. *Mol. Cell* 49, 439–452.
- Siegel, V., and Walter, P. (1988). Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* 52, 39–49.
- Stefani, G., Fraser, C.E., Darnell, J.C., and Darnell, R.B. (2004). Fragile X mental retardation protein is associated with translating polyribosomes in neuronal cells. *J. Neurosci.* 24, 7272–7276.
- Sutton, M.A., Taylor, A.M., Ito, H.T., Pham, A., and Schuman, E.M. (2007). Postsynaptic decoding of neural activity: eEF2 as a biochemical sensor coupling miniature synaptic transmission to local protein synthesis. *Neuron* 55, 648–661.
- Takyar, S., Hickerson, R.P., and Noller, H.F. (2005). mRNA helicase activity of the ribosome. *Cell* 120, 49–58.
- Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344–354.
- Udagawa, T., Farny, N.G., Jakovcevski, M., Kaphzan, H., Alarcon, J.M., Anilkumar, S., Ivshina, M., Hurt, J.A., Nagaoka, K., Nalavadi, V.C., et al. (2013). Genetic and acute CPEB1 depletion ameliorate fragile X pathophysiology. *Nat. Med.* 19, 1473–1477.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.
- Zaborske, J.M., DuMont, V.L.B., Wallace, E.W.J., Pan, T., Aquadro, C.F., and Drummond, D.A. (2014). A nutrient-driven tRNA modification alters translational fidelity and genome-wide protein coding across an animal genus. *PLoS Biol.* 12, e1002015.
- Zhang, Y., Ma, C., Yuan, Y., Zhu, J., Li, N., Chen, C., Wu, S., Yu, L., Lei, J., and Gao, N. (2014). Structural basis for interaction of a cotranslational chaperone with the eukaryotic ribosome. *Nat. Struct. Mol. Biol.* 21, 1042–1046.

Multisensory Signaling Shapes Vestibulo-Motor Circuit Specificity

Graphical Abstract



Authors

Emanuela Basaldella, Aya Takeoka, Markus Sigrist, Silvia Arber

Correspondence

silvia.arber@unibas.ch

In Brief

Vestibular and proprioceptive system signaling controls body balance by targeting motor neurons regulating functionally distinct muscles and fiber types. Their synaptic inputs at the final step of motor output processing are shaped by functional interactions between these two sensory systems.

Highlights

- Preferential targeting of vestibular input to slow motor neurons of extensor pools
- Vestibular input specificity to motor neurons assembled by developmental refinement
- Vestibular and proprioceptive input to motor neuron subtypes inversely correlated
- Proprioceptive circuits shape vestibular input to specific motor neuron subtypes



Multisensory Signaling Shapes Vestibulo-Motor Circuit Specificity

Emanuela Basaldella,^{1,2} Aya Takeoka,^{1,2} Markus Sigrist,^{1,2} and Silvia Arber^{1,2,*}

¹Biozentrum, Department of Cell Biology, University of Basel, 4056 Basel, Switzerland

²Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland

*Correspondence: silvia.arber@unibas.ch

<http://dx.doi.org/10.1016/j.cell.2015.09.023>

SUMMARY

The ability to continuously adjust posture and balance is necessary for reliable motor behavior. Vestibular and proprioceptive systems influence postural adjustments during movement by signaling functionally complementary sensory information. Using viral tracing and mouse genetics, we reveal two patterns of synaptic specificity between brainstem vestibular neurons and spinal motor neurons, established through distinct mechanisms. First, vestibular input targets preferentially extensor over flexor motor pools, a pattern established by developmental refinement in part controlled by vestibular signaling. Second, vestibular input targets slow-twitch over fast motor neuron subtypes within extensor pools, while proprioceptors exhibit inversely correlated connectivity profiles. Genetic manipulations affecting the functionality of proprioceptive feedback circuits lead to adjustments in vestibular input to motor neuron subtypes counterbalancing the imposed changes, without changing the sparse vestibular input to flexor pools. Thus, two sensory signaling systems interact to establish complementary synaptic input patterns to the final site of motor output processing.

INTRODUCTION

Descending motor control pathways are essential to regulate spinal circuits involved in movement (Grillner and Dubuc, 1988; Lundberg, 1975). Specificity of synaptic connections between upper motor control centers and the spinal output system provides the anatomical substrate to implement movement variety and precision. As animals grow up, they engage in progressively more diverse and refined motor behaviors, paralleling the establishment of functionally mature descending input to spinal circuits. Despite the importance of this descending connection matrix, the organization of its key components and especially the elucidation of developmental mechanisms involved in its establishment are still under intense investigation.

The ability to continuously adjust posture and balance during movement matures at postnatal stages in mammals (Brown,

1981; Geisler et al., 1993). Due to the importance of these adaptive mechanisms for the execution of highly diverse motor programs, circuits steering body stabilization must exhibit a high degree of tuning flexibility. Two parallel and functionally complementary sensory signaling systems play key roles in this process. In the vestibular system, one central sensory organ in the inner ear monitors linear and rotational acceleration and provides input to the vestibular nucleus of the brainstem (Angelaki and Cullen, 2008; Brodal and Pompeiano, 1957). Descending vestibulo-spinal projection neurons transmit this information to spinal circuits to provide postural stability (Grillner et al., 1970; Lund and Pompeiano, 1968; Shinoda et al., 1988; Wilson and Yoshida, 1968). The somatosensory system represents a complementary signaling system in which sense organs are distributed throughout the entire body (Abraira and Ginty, 2013; Brown, 1981; Matthews, 1981). Within this system, proprioceptive sensory neurons located in dorsal root ganglia (DRG) monitor self-generated actions and extrinsic perturbations in the periphery. Of these, muscle spindle afferents report the state of muscle contraction from specific sites in the periphery directly to spinal motor neurons through monosynaptic reflex arcs (Brown, 1981; Eccles et al., 1957; Windhorst, 2007).

Revealing the organization of synaptic connections to spinal motor neurons is crucial to understand how vestibular and proprioceptive information influences motor output. Studies in the adult cat provide the first evidence that vestibular neurons preferentially target extensor motor neuron pools (Grillner et al., 1970). In contrast, proprioceptors contact motor neurons of most pools in the spinal cord. A motor pool receives direct synaptic input from muscle spindle afferents supplying the same or synergistic muscles, but not from afferents innervating antagonistic muscles (Eccles et al., 1957; Mears and Frank, 1997). Thus, both extensor and flexor motor neuron pools get direct proprioceptive input but in highly specific configurations, whereas direct vestibular input seems to be preferentially targeted to extensor motor neurons in line with its body-stabilizing and anti-gravitational function. Beyond their connectivity profiles, vestibular and proprioceptive systems also interact functionally with each other and can contribute to both enhancement or depression of responses in motor neurons (Grillner et al., 1970).

Less is known about the mechanisms guiding developmental assembly of these two sensory systems. Specific connectivity between proprioceptors and motor neurons in the same reflex arc is already present at early postnatal developmental stages in mice (Mears and Frank, 1997) and activity-independent in

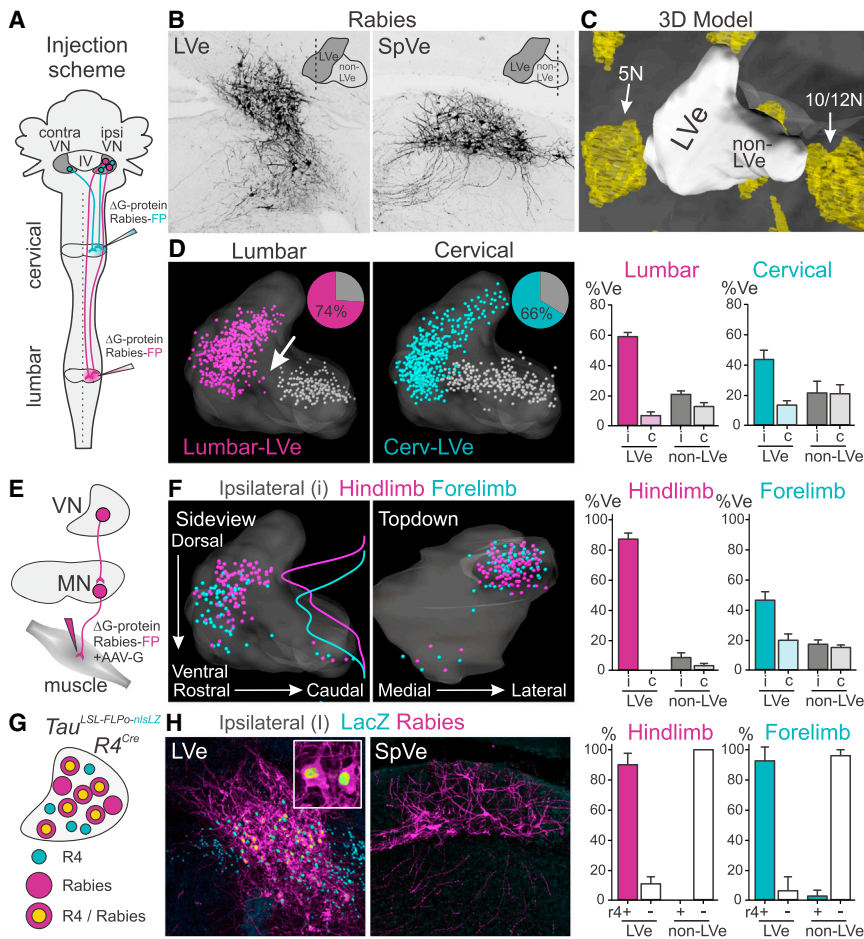


Figure 1. Spatial Distribution of Vestibular Neurons Regulating Spinal Motor Neurons

(A) Unilateral injections of G-protein-deleted Rabies viruses encoding fluorescent proteins (FP) into cervical and lumbar spinal cord to assess the position of vestibular neurons in the brainstem.

(B) Representative coronal sections at the level of the lateral vestibular nucleus (LVe; left) and spinal vestibular nucleus (SpVe; right) ipsilateral to injection.

(C) Three-dimensional model of the vestibular nucleus used for digital reconstructions, surrounded by cranial motor nuclei 5N and 10/12N.

(D) Ipsilateral side view of digital 3D vestibular nucleus reconstructions derived from lumbar (left) and cervical (middle) spinal injections (colored neurons reside in, gray neurons outside LVe). Pie charts in upper right corners show percentages of LVe neurons. Right: quantification of lumbar and cervical projection neuron composition in Ve nucleus, stratified by ipsi- and contralateral as well as LVe and non-LVe residence.

(E) Strategy for monosynaptic rabies tracing experiments to determine connectivity between vestibular and motor neurons.

(F) Side- (left) and top-down (middle) view of vestibular nucleus ipsilateral to muscle injection, depicting the position of vestibular neurons connected to FL- (cyan) and HL- (purple) innervating motor neurons. Density curves for HL and FL premotor neuron distributions along the dorso-ventral axis superimposed to the ipsilateral side-view panel. Right: quantification of positional distribution as in (D).

(G) Genetic strategy to mark LVe neurons by developmental origin. *R4::Cre* mice are crossed to *Tau*-reporter mice for conditional

expression of nls-LacZ and FLPo expression to assess the percentage of premotor vestibular neurons marked by R4-origin.

(H) Most HL- or FL-premotor (Rabies^{DN}) neurons in LVe are marked by *R4::Cre*-induced LacZ, whereas premotor neurons residing in SpVe do not carry this tag (left, middle: exemplary images; right: quantification).

frogs (Frank, 1990). In the vestibular system, transient developmental perturbations affect motor behavior in several species (Geisler and Gramsbergen, 1998; Moorman et al., 2002; Van Cleave and Shall, 2006; Walton et al., 2005), raising the possibility that the assembly of the vestibular system might be plastic. Together, these observations provide first hints that even though proprioceptive and vestibular systems both functionally converge on motor neurons, their organization and developmental assembly mechanisms might be distinct. Moreover, whether and how they influence each other to establish mature functionality is unknown.

In this study, we exploit intersectional viral tracing technology and mouse genetics to reveal that vestibulo-spinal projection neurons in the brainstem exhibit connection specificity to motor neurons. We demonstrate that they do not only target extensor over flexor motor neuron pools, but that within extensor pools, they preferentially connect to slow over fast motor neurons. We find that connectivity profiles arise gradually at postnatal developmental stages, paralleling postural maturation. Genetic perturbation of vestibular signaling leads to interpool connecti-

vity defects, whereas proprioceptive feedback circuit alterations induce specific connectivity shifts in synaptic scaling of vestibular input to motor neuron subtypes. These findings support a model in which two major sensory signaling systems interact at the final motor output step to establish specific connectivity profiles by complementary cross-modal signaling.

RESULTS

Spatial Organization of Spinal Projection Neurons in the Vestibular Nucleus

To delineate the position of vestibular neurons with spinal projections, we performed unilateral intraspinal injections of G-protein-deficient rabies viruses encoding fluorescent marker proteins (FP) (Rab-FP) (Wickersham et al., 2007) (Figures 1A and 1B). We found that in a three-dimensional digital brainstem model (Figure 1C), vestibular neurons with lumbar projections were preferentially located ipsilaterally, with dominant residence within the lateral vestibular (LVe) nucleus and with a clear spatial segregation to a caudal cluster of non-LVe neurons that were

bilaterally distributed (Figure 1D; including spinal vestibular neurons [SpVe]). Vestibular neurons projecting to cervical spinal levels also showed clear, albeit less pronounced ipsilateral residence within LVe, but occupied the vestibular nucleus continuously into caudal non-LVe territory (Figure 1D). In summary, for both lumbar and cervical vestibular projection neurons, LVe neurons exhibited a strong ipsilateral bias (Figure 1D). These findings confirm and extend findings that subgroups of mouse vestibular neurons exhibit differential projection trajectories to interact with local circuits in the spinal cord (Liang et al., 2014, 2015).

To determine the identity of vestibular neurons exerting the most direct influence on spinal motor neurons, we next assessed abundance and position of vestibular neurons with direct synaptic connections to motor neurons. We used a transsynaptic rabies virus-based approach with monosynaptic restriction (Wickersham et al., 2007) (Figure 1E). The majority of neurons with direct connections to lumbar motor neurons resided in the ipsilateral LVe nucleus, with the highest density peak more dorsal to neurons with connections to cervical motor neurons (Figure 1F).

To gain genetic access to neurons in the LVe nucleus, we applied a lineage tracing approach for neurons developmentally derived from different rhombomeric (R) origin (Figure 1G). To permanently mark R4-derived neurons, we used intersectional breeding of *R4::Cre* mice (Di Bonito et al., 2013) and the conditional neuronal reporter strain *Tau^{lox-STOP-lox-Flp-INLA}* (Pivet et al., 2014). This strategy labeled the majority of lumbar-projecting LVe neurons, but the R4-marker was entirely excluded from non-LVe neurons (Figure 1H), which are derived from more caudal rhombomeres (data not shown). Together, the existence of the clearly delineated ipsilateral cluster of vestibular neurons in the LVe nucleus and the access to specific targeting approaches allowed us to next dissect projection trajectory and connection specificity of these neurons to lumbar spinal circuits with precision.

Lateral Vestibular Synaptic Input Is Biased to Extensor Motor Neurons

To reveal the descending projection trajectory and the synaptic arborization pattern of LVe neurons to the lumbar spinal cord, we performed focal injections of adeno-associated viruses (AAV) into the LVe nucleus. We used AAVs expressing tdTomato for axonal tracing and/or a fusion protein between synaptophysin and GFP or Myc (Syn-Tag) for synaptic reconstructions (Figures S1A–S1C). We found that axons descending from the LVe nucleus to the lumbar spinal cord were confined to ipsilateral white matter tracts (Figure S1D), consistent with previous experiments (Liang et al., 2014). Analysis of Syn-Tag distribution in the lumbar spinal cord revealed the highest density of synaptic terminals in lamina VIII ipsilateral to injection (Figures S1D and S2). Many Syn-Tag puncta were also detected throughout the ipsilateral ventral spinal cord below the central canal including lamina IX containing ChAT^{ON} motor neurons (Figures S1D and S2). A similar distribution pattern was observed upon AAV-FRT-Syn-Tag LVe injection in *R4::Cre/Tau^{lox-STOP-lox-Flp-INLA}* mice (Figure S2). Moreover, and similar to findings in the rat (Du Beau et al., 2012), the majority of Syn-Tag^{ON} terminals accumulate the vesicular glutamate

transporter vGlut2 (74.6%; Figure S1E), demonstrating that LVe spinal projection neurons provide excitatory input to the lumbar spinal cord.

We next assessed whether LVe input to the lumbar spinal cord exhibits synaptic specificity with respect to the identity of contacted motor neurons. We combined LVe AAV-Syn-Tag injections with retrograde tracing of motor neurons from identified hindlimb muscles (Figures 2A and S1A). We analyzed LVe input to motor neurons pools innervating the ankle extensor gastrocnemius (GS) and the ankle flexor tibialis anterior (TA), due to their functional antagonism as well as previous evidence for GS-biased vestibular synaptic input in the cat (Grillner et al., 1970). Vestibular input was also strongly biased toward the GS compared to the TA motor neuron pool in mice, a bias detected irrespective of cell body or dendritic analysis of reconstructed GS/TA motor neurons (Figures 2A and 2B).

Lateral Vestibular Input Avoids GSL1 Motor Neuron Subtypes

Despite this striking difference in overall input between GS and TA motor neurons, we noted that LVe synaptic input to individual GS motor neurons was highly variable. While some GS motor neurons received low (GS-low) LVe input, others were targeted by high-density (GS-high) LVe input (Figures 2A and 2B). These findings suggest that not all GS motor neurons are equally favored targets for vestibular input and raise the question of the underlying reason for this variability.

Most skeletal muscles are composed of a mixture of different fiber types innervated by three functionally matched alpha motor neuron subpopulations. These motor neuron subtypes are differentially recruited during movement and include fast fatigable (FF), fatigue resistant (FR), and slow motor units (Burke, 1967; Kanning et al., 2010). In the mouse, the most lateral subcompartment of the lateral GS muscle (GSL1) is a very valuable exception to this rule in that it is innervated exclusively by FF motor neurons (Pun et al., 2006). This property allowed us to assess LVe input specifically to FF motor neurons within the GS motor pool (Figure 2C). We found that GSL1 FF motor neurons received only low-density LVe input and notably significantly less than the entire GS motor pool (Figures 2C and 2D). In addition, cell body volume values of motor neurons innervating the GSL1 compartment have a tendency to accumulate in the upper two-thirds of the distribution spectrum (Figure 2E). Nonetheless, and consistent with previous observations (Burke et al., 1982), such size range classifications are not sufficient to unambiguously assign motor neuron subtype identity. In summary, GSL1 FF motor neurons receive low-density LVe input, raising the possibility that this input is preferentially targeted to specific motor neuron subtypes within extensor pools.

LVe Input Prefers Molecularly Defined Slow Motor Neurons in Extensor Pools

We next aimed to generalize our finding that LVe inputs might prefer slow motor neuron subtypes. Recent observations demonstrate that chondroitin (Chodl) and matrix metalloproteinase-9 (MMP-9) are expressed by fast motor neurons (Enjin et al., 2010; Kaplan et al., 2014; Leroy et al., 2014). In mice expressing the membrane marker protein placental alkaline

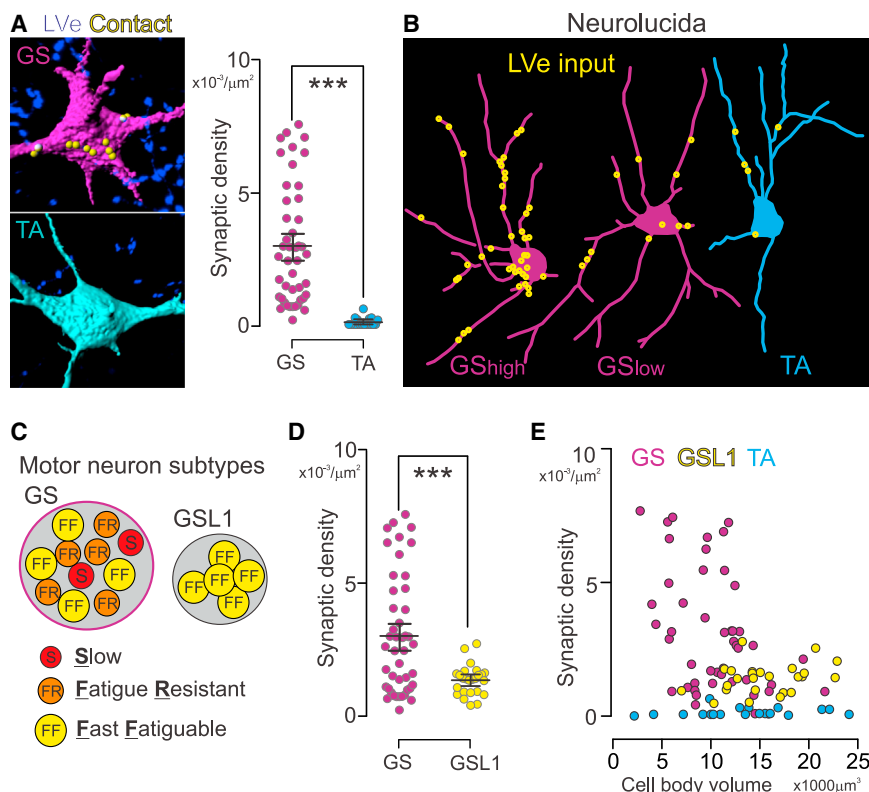


Figure 2. Vestibular Input Stratifies by Motor Neuron Subtype and Size

(A) Digital reconstruction and quantification of LVe synaptic input density to GS and TA motor neurons (each dot represents one motor neuron).

(B) Representative Neurolucida reconstructions of GS/TA motor neurons and LVe synaptic input (yellow). GS examples with high and low input density are shown.

(C) Motor neuron subtype composition of GS and GSL1 motor pool stratified into slow (S), FR (fatigue resistant), and FF (fast fatiguable) subtypes.

(D) Quantification of synaptic density of LVe input to GS and GSL1 motor neurons (each dot represents one motor neuron).

(E) Synaptic density of LVe input to analyzed motor neurons plotted against cell body volumes. See also Figures S1 and S2.

phosphatase (PLAP) from the *Chodl* locus (*Chodl*^{PLAP}) (Sakurai et al., 2013), a large majority of *Chodl*^{ON} lumbar ChAT^{ON} motor neurons in the lateral motor column (LMC) coexpressed MMP-9 (92%), and all GSL1 FF motor neurons were PLAP^{ON}/MMP-9^{ON} (Figures 3A–3C). We first quantified LVe synaptic input density to lumbar LMC motor neurons overall in *Chodl*^{PLAP} mice. Stratification of LVe input density by PLAP^{ON} and PLAP^{OFF} status of targeted LMC motor neurons revealed significantly lower input density to PLAP^{ON} than putative alpha PLAP^{OFF} motor neurons (Figures 3D and 3E). Furthermore, there was a significant inverse correlation between LVe synaptic input density and motor neuron cell body volume (Figure 3E).

To determine whether the uncovered LVe synaptic input rule based on *Chodl*/MMP-9 stratification also applies to motor neuron subtypes within a given extensor motor pool other than GS (Figures 2C–2E), we analyzed two more motor pools innervating extensor muscles. The ankle extensor muscle soleus (Sol) is innervated by an approximately equal number of slow and FR motor neurons in mice, but does not contain any FF motor neurons (Kaplan et al., 2014; Pun et al., 2006). Analysis of LVe input density to Sol motor neurons stratified by MMP-9 status revealed significantly lower values for MMP-9^{ON} than MMP-9^{OFF} Sol motor neurons (Figure 3F). Additionally, since the Sol motor pool does not contain any FF motor neurons (Pun et al., 2006), these findings indicate that slow motor neurons are not only a preferred LVe target over FF, but also over FR motor neurons. We next assessed LVe input to MMP-9 stratified motor neurons innervating the hip extensor gluteus (GL) and found that also for this pool, MMP-9^{ON} populations received significantly lower

input than the MMP-9^{OFF} cohort (Figure 3F). Moreover, the corresponding functionally antagonistic hip flexor (Iliopsoas) motor pool showed LVe input density values similarly low as to TA flexor motor neurons (Figure 3F), thus generalizing our findings to other motor neuron pools.

Together, our experiments support a model in which synaptic input specificity

of LVe neurons to lumbar LMC motor neurons is organized at different levels (Figure 3G). First, LVe axons seek out extensor over flexor motor pools as preferred synaptic targets in agreement with previous work (Grillner et al., 1970). Second, LVe synaptic contacts preferentially target slow over fast motor neuron subtypes within an extensor pool. These findings raise the question of how this synaptic specificity arises during development and what may be factors regulating its establishment.

Developmental Refinement of Vestibular Synaptic Input Specificity to Motor Neurons

To assess synaptic input specificity of LVe neurons to lumbar motor neurons during development, we carried out spatially confined injections of AAV-Syn-Tag into the LVe nucleus early postnatally and retrogradely labeled GS or TA motor neurons (Figure 4A). The earliest time point for which it was technically possible to achieve consistent high-level Syn-Tag accumulation from LVe neurons in the lumbar spinal cord was P7. GS motor neurons at P7 receive synaptic input at densities similar to adult (Figure 4B). However, while the difference in input density between GS and TA motor neurons was already established at P7, LVe terminals frequently contacted TA motor neurons at an overall significantly higher input density than in the adult (Figure 4B). To assess during which time window the transition to mature connectivity profiles emerges, we carried out synaptic input mapping at progressively later developmental time points (Figure 4A). We found that LVe neurons still contact TA motor neurons at P11, but that developmental refinement was complete by P17 (Figure 4B).

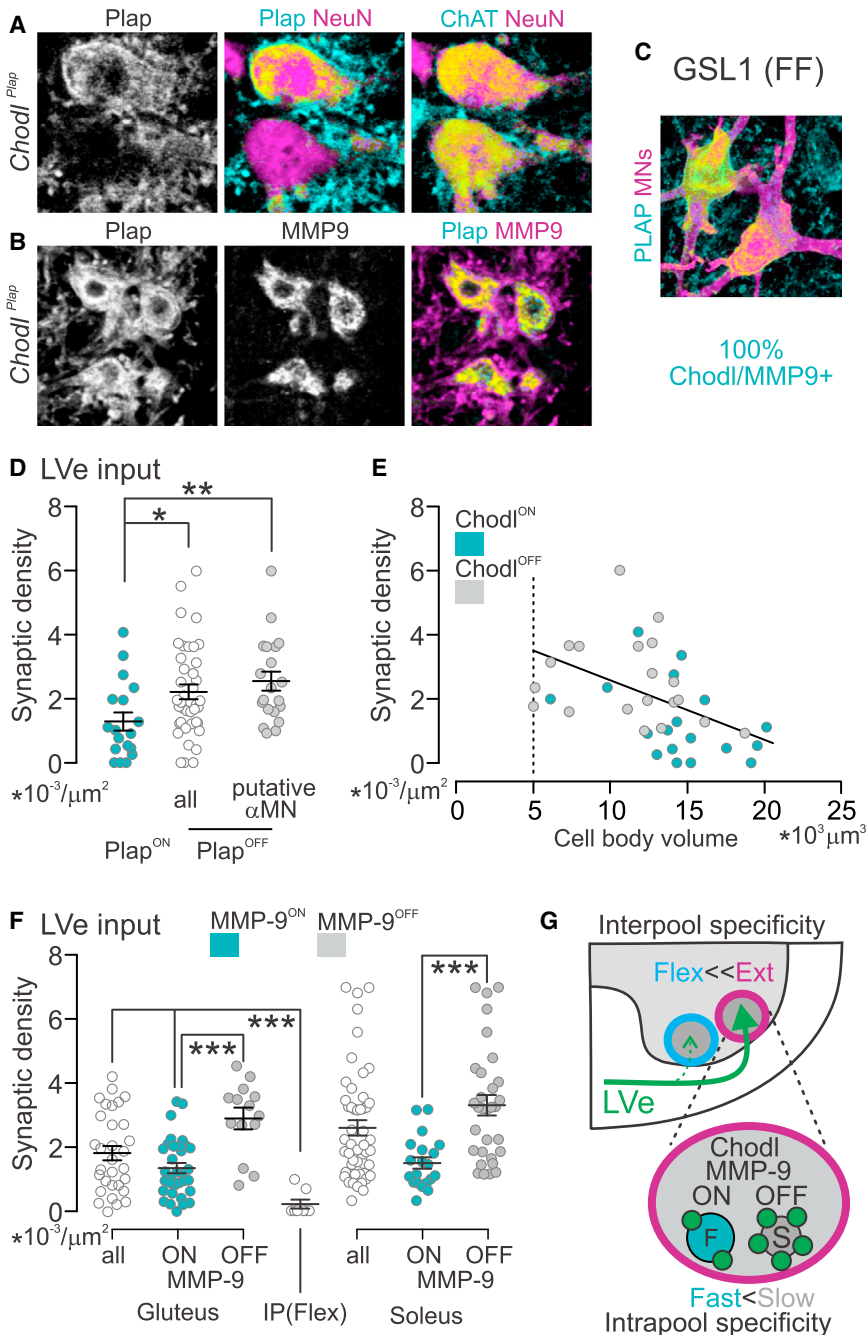


Figure 3. Vestibular Input Preferentially Targets Putative Slow over Fast Motor Neurons

(A and B) ChAT^{ON}/NeuN^{ON} alpha motor neurons in the lumbar spinal cord of *Chodl*^{PLAP} mice fractionate into Chodl^{ON} and Chodl^{OFF} population, of which the Chodl^{ON} neurons also express MMP-9. (C) Retrogradely marked GSL1 FF motor neurons express Chodl.

(D) Density of Lve synaptic input to PLAP^{ON}, PLAP^{OFF} (all, or excluding putative gamma motor neurons using a size cut-off criterion of 5,000 μm^3) motor neurons.

(E) Synaptic density of Lve input to PLAP^{ON} and putative alpha PLAP^{OFF} motor neurons analyzed in (D) plotted against cell body volumes ($r = -0.548$, $p = 0.0005$).

(F) Analysis of Lve input density for gluteus and soleus motor neurons stratified by MMP-9 expression status (cyan, MMP-9^{ON}; gray, MMP-9^{OFF}). Input to antagonistic hip flexor muscle iliopsoas (IP) is also shown.

(G) Summary diagram of synaptic specificity between Lve and motor neurons. Lve preferentially targets extensor over flexor motor pools (top, interpool specificity) and within extensor pools preferentially slow over fast motor neuron subtypes (bottom, intrapool specificity, green dots represent synapses).

neurons are eliminated between P11 and P17, when they reach a mature connectivity profile (Figure 4F).

Perturbing Vestibular Signaling Affects Establishment of Interpool Synaptic Specificity

To elucidate the mechanisms by which selectivity of vestibular input to spinal motor neurons is established, we used two different genetic models in the mouse exhibiting altered vestibular neuron signaling. We asked how these perturbations influence the establishment of mature connectivity profiles between vestibular neurons and spinal motor neurons.

We first analyzed *NADPH oxidase 3* (*Nox3*) mutant mice (Figure 5A). These mice lack mineralized particles called otoconia in the inner ear's utricle and

To determine whether Lve contacts to TA motor neurons at early postnatal stages represent synaptic contacts, we applied monosynaptic rabies viruses to muscles innervated by GS and TA motor neurons (Figure 4C). We found that Lve neurons connect to both GS and TA motor pools at these stages, but significantly more Lve neurons were labeled after GS than TA muscle injections, at a ratio comparable to the anterograde synaptic density measurements at P11 (Figures 4D and 4E). Together, these data confirm our anterograde tracing results, demonstrating that initial developmental synaptic contacts to TA motor

sacculle, leading to selective defects in perception of gravity and linear acceleration, but they exhibit intact semicircular canal vestibular as well as auditory sensory inputs (Paffenholz et al., 2004). Of the five known vestibular input channels, predominantly utricular or posterior semicircular canal nerve activation influences lumbar spinal circuits through the lateral vestibular tract (Uchino and Kushiro, 2011). *Nox3* mutant mice therefore exhibit congenitally altered Lve input to the lumbar spinal cord, lacking information derived from the utricular sensory input channel but not from semicircular canals.

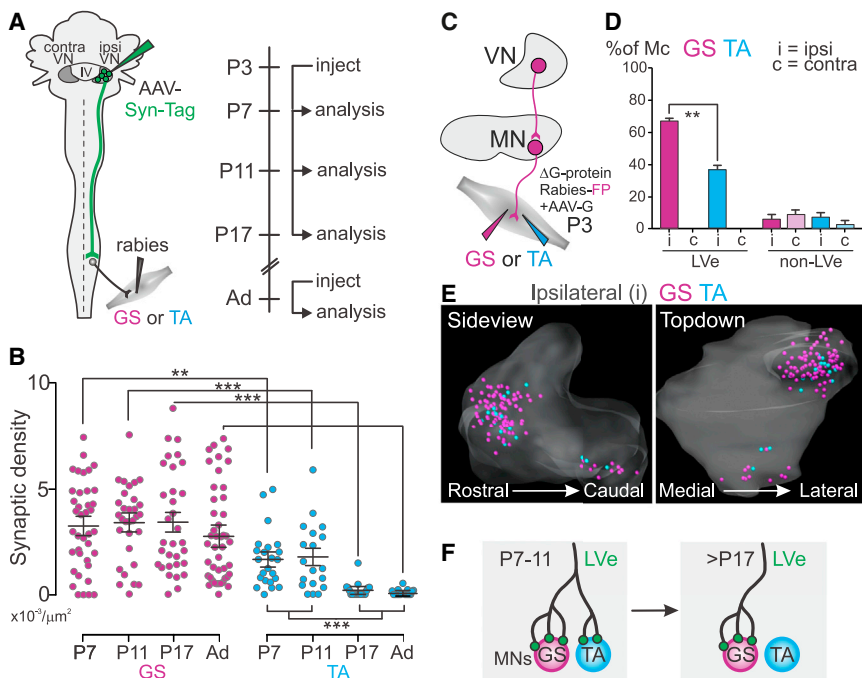


Figure 4. Developmental Refinement of Vestibular Input to Lumbar Motor Neurons

(A) Experimental approach used and timeline. AAV-Syn-Tag is injected into LVe of P3 (or adult) mice, followed by retrograde marking of GS or TA motor neurons by muscular tracer injections. (B) Synaptic density of LVe input to GS and TA motor neurons at P7, P11, P17, and adult stages. (C–E) Monosynaptic rabies tracing experiment at P3. Quantification of marked neurons in LVe and non-LVe territory (see Figure 1), both ipsi- (i) and contra- (c) lateral to muscle injection (normalized to rabies neuron number in Magnocellular nucleus). Side- and top-down view of ipsilateral vestibular reconstruction depicting GS (purple) and TA (cyan) vestibular neurons shown in (E). Neurons connected to GS or TA motor neurons were intermingled and no spatial segregation was discernible (E). (F) Summary diagram illustrating developmental refinement process of LVe input to GS and TA motor neurons.

We first determined whether *Nox3* mutation affects the establishment of LVe synaptic inputs to the functionally antagonistic motor neuron pools GS and TA. We found that there was no difference in LVe input density to GS motor neurons between wild-type and *Nox3* mutant mice, but that TA motor neurons received LVe input at a significantly higher density in *Nox3* mutant than wild-type mice (Figure 5B). When we compared LVe synaptic input density at P11, a time point before mature connectivity profiles are reached in wild-type mice, LVe input to TA motor neurons was not different between wild-type and *Nox3* mutant mice (Figure 5B). Moreover, between P11 and adult stages in *Nox3* mutant mice, no significant refinement of LVe input to TA motor neurons occurred (Figure 5B). Together, these findings demonstrate that LVe neurons maintain aberrant synaptic input to flexor motor neurons when otolithic vestibular signaling is non-functional.

We next asked whether utricular vestibular signaling also influences LVe connectivity profiles to fast and slow motor neuron subtypes. There was no significant difference in LVe synaptic input to FF GSL1 motor neurons between wild-type and *Nox3* mutant mice (Figure 5C). We also analyzed LVe input density to Sol motor neurons stratified by MMP-9 status to distinguish between fast (MMP-9^{ON}) and slow (MMP-9^{OFF}) motor neuron subtypes. While *Nox3* mutant mice still exhibited clear intrapool differences to these motor neuron subtypes, the connectivity stratification was less pronounced than in wild-type mice (Figure 5C). Together, these findings suggest that *Nox3* mutants exhibit defects in interpool but no major intrapool LVe synaptic connectivity.

We next analyzed an intersectional mouse mutant in which the synaptic output of most LVe neurons is functionally muted from the earliest developmental stages. This genetic strategy is based on our observations that most LVe neurons projecting to lumbar

spinal levels are of developmental rhombomeric origin R4 and express the glutamate transporter vGlut2. In agreement, genetic elimination of vGlut2 from R4-derived LVe neurons (*R4^{Cre}::vGlut2^{fllox}* mice) abolishes vGlut2 protein from the vast majority of spinal synapses derived from LVe neurons (Figures 5A and S3).

Since *R4^{Cre}::vGlut2^{fllox}* mice have not been characterized before, we determined whether they exhibit motor behavioral deficiencies compatible with impaired LVe function. *R4^{Cre}::vGlut2^{fllox}* mice executed open field navigation, grip strength, and horizontal ladder tasks similar to wild-type mice (Figure 5D). In contrast, they exhibited defects in tasks predicted to profoundly engage the vestibular system. *R4^{Cre}::vGlut2^{fllox}* mice walking on a narrow beam showed significantly more slips than wild-type mice, and this phenotype was particularly pronounced on 6 mm- over 12 mm-wide beams (Figure 5D). These behavioral experiments suggest that elimination of vGlut2 from R4-derived LVe neurons affects vestibular function and leads to motor defects attributable to such perturbations.

We next assessed synaptic input to TA motor neurons in these mice and found a significantly higher synaptic input density compared to wild-type (Figure 5B), similar to the phenotype in *Nox3* mutant mice. Lastly, also similar to our observations in *Nox3* mutant mice, we found no differences in LVe synaptic input to GSL1 and Sol motor pools in *R4^{Cre}::vGlut2^{fllox}* compared to wild-type mice (Figure 5C).

In summary, genetic perturbation of selective vestibular input channels or muting synaptic output of vestibular neurons result in similar connectivity defects between LVe neurons and flexor motor neurons (Figure 5E). Our observations also reveal that additional factors must play important roles in scaling vestibular input specificity to motor neuron subtypes. Considering the established roles of vestibular and proprioceptive systems in posture and balance, an interesting hypothesis to test is whether these two systems influence each other in establishing their respective connection specificities to motor neurons.

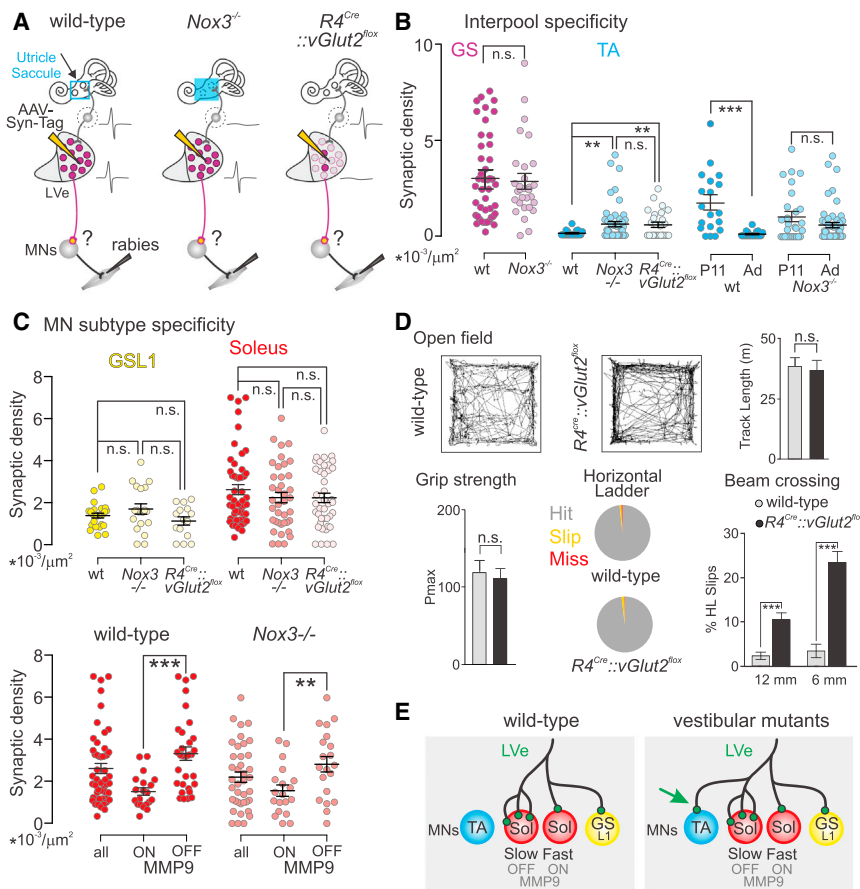


Figure 5. Perturbation of Vestibular Input Channel Results in Connectivity Defects to Motor Neurons

(A) Cellular phenotypes and analysis of wild-type, $Nox3^{-/-}$, and $R4^{Cre}::vGlut2^{lox}$ mice. $Nox3^{-/-}$ mice exhibit defects in otolith-organ derived vestibular sensory input to brainstem vestibular neurons, whereas $R4^{Cre}::vGlut2^{lox}$ mice lack functional output from vestibular neurons to the spinal cord. AAV-Syn-Tag injections are performed to quantify synaptic input density to motor neuron subpopulations.

(B) Synaptic density of LVe input to GS and TA motor neurons at adult (GS, TA) and P11 (TA) stages for wild-type and $Nox3$ mutant mice. TA motor neurons were analyzed in adult $R4^{Cre}::vGlut2^{lox}$ mice.

(C) Synaptic density of LVe input to GSL1 and Sol motor neurons in wild-type, $Nox3$ mutant, and $R4^{Cre}::vGlut2^{lox}$ mice (top row). Data for Sol motor neurons in wild-type and $Nox3$ mutant mice displayed stratified by MMP-9 expression status (bottom row).

(D) Behavioral analysis of wild-type and $R4^{Cre}::vGlut2^{lox}$ mice in open field arena (tracks of individual mice; quantification of track length moved in 10 min), grip strength, horizontal ladder precision (hit, slip, and miss categories displayed in pie chart) and beam crossing on 12 mm- and 6 mm-thick beam.

(E) Summary diagram of synaptic input analyzed between LVe and motor neuron subtypes in wild-type mice and vestibular mutants. Note ectopic synaptic input to TA motor neurons in vestibular mutants.

See also Figure S3.

Proprioceptive Signaling Influences Vestibular Synaptic Density to Motor Neurons

Given the striking LVe synaptic input variation to different motor neuron subtypes, we first determined the organization of direct synaptic input by proprioceptive afferents to motor neuron subtypes. Of proprioceptors, only muscle spindle afferents connect directly to motor neurons and their synaptic terminals accumulate the vesicular glutamate transporter vGlut1 (Oliveira et al., 2003; Pecho-Vrieseling et al., 2009). Analogous to our analysis of LVe input to motor neuron subtypes (Figure 3), we quantified vGlut1 input density to $Chodl^{ON}$ putative fast motor neurons, $Chodl^{OFF}$ putative slow motor neurons, as well as to identified GSL1 (exclusively FF) and Sol (many slow) motor neurons. We found that vGlut1 input density was higher for GSL1 and $Chodl^{ON}$ motor neurons than for Sol and $Chodl^{OFF}$ putative alpha motor neurons (Figure 6A), a finding opposite to our analysis of input densities derived from the LVe nucleus (Figures 3D and 3E). Moreover, cell body volumes and vGlut1 synaptic input density were positively correlated to each other (Figure 6B), further supporting the notion that fast motor neurons with relatively large cell bodies receive a higher density of vGlut1 inputs than smaller, $Chodl^{OFF}$ alpha motor neurons. Analysis of both LVe and vGlut1 input to the same cohort of motor neurons stratified by $Chodl$ -expression status and cell size confirmed this conclusion (Figures 6C and 6D).

To determine whether the status of proprioceptive input to a motor neuron influences the organization of LVe input to the same motor neuron, we analyzed two mouse mutants with opposite proprioceptive synaptic phenotypes to motor neurons (Figure 7A). *Egr3* mutant mice exhibit early postnatal degeneration of muscle spindles, leading to non-functional muscle spindle afferents (Chen et al., 2002; Tourtellotte and Milbrandt, 1998). In contrast, *Mic::NT3* mice overexpress NT3 from skeletal muscle fibers, resulting in survival of superfluous proprioceptive afferents with aberrant and more synaptic connections to central synaptic partners (Wang et al., 2007). To assess LVe input to motor neurons in these two mutant mouse strains compared to wild-type mice, we quantified synaptic input to motor neuron cell bodies and dendrites (Figures 7A and 7B).

In *Egr3* mutant mice, we analyzed LVe and vGlut1 input to GSL1 motor neurons, normally exhibiting high-vGlut1 and low-LVe input (Figures 7C–7E). In these mice, vGlut1 contacts to GSL1 motor neurons are present (Figures 7D and 7E) despite their non-functionality (Chen et al., 2002; Tourtellotte and Milbrandt, 1998). However, LVe input to GSL1 motor neurons is significantly increased in *Egr3* mutant mice (Figures 7D and 7E). Conversely, in *Mic::NT3* mice, we analyzed LVe and vGlut1 input to Sol motor neurons that normally receive relatively low-vGlut1 and high-LVe input (Figures 7F–7H). As expected, Sol motor neurons received significantly more vGlut1

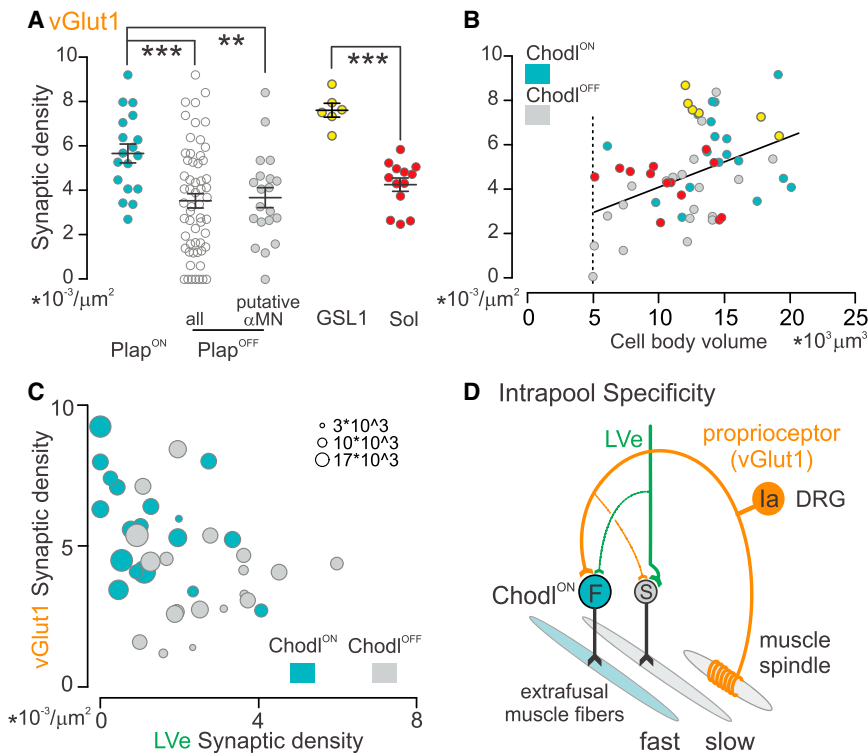


Figure 6. Vestibular and Proprioceptive Input Anti-correlated by Motor Neuron Subtype

(A) Density of vGlut1 synaptic input to PLAP^{ON}, PLAP^{OFF} (all, or excluding putative gamma motor neurons using a size cut-off criterion of $5,000 \mu\text{m}^3$), GSL1, and soleus motor neurons.

(B) Synaptic density of vGlut1 input to PLAP^{ON} and putative alpha PLAP^{OFF} motor neurons analyzed in (A) plotted against cell body volumes ($r = 0.448$, $p = 0.0054$). GSL1 and Sol motor neurons are also displayed in this plot but not included in the correlation analysis.

(C) Plot of vGlut1 versus LVe synaptic input density to PLAP^{ON} and putative alpha PLAP^{OFF} motor neurons in relation to cell body volume illustrated by diameter of plotted circles.

(D) Intrapool stratification of LVe and Ia proprioceptive vGlut1 input to fast (F, Chodl^{ON}) and slow (S, Chodl^{OFF}) alpha motor neurons, revealing anti-correlated synaptic input densities.

input in *MLC::NT3* than wild-type mice, but LVe input density was strongly reduced (Figures 7G and 7H). Despite these differences in LVe connectivity to motor neuron subtypes however, overall LVe synaptic patterns in the spinal cord were not perturbed across genotypes and injection conditions (Figure S4).

To determine whether the lack of direct functional proprioceptive input to motor neurons in *Egr3* mutant mice also influences LVe input to flexor motor neurons, we next compared input to TA motor neurons between wild-type and *Egr3* mutant mice. We found that there was no significant difference between genotypes (Figures S5A–S5D). These results demonstrate that altered proprioceptive signaling to flexor motor neurons cannot overrule the scarcity of LVe input to these neurons. Thus, the assembly of LVe inputs at the motor pool and motor neuron subtype level employs distinct developmental mechanisms.

Lastly, to test whether the synaptic scaling of these two complementary sensory systems operates bidirectionally, i.e., whether altered LVe input scales vGlut1 input to motor neurons, we analyzed vGlut1 input to Sol motor neurons in *Nox3* mutant mice. We detected a striking increase in vGlut1 terminals to Sol motor neurons in these mutants compared to wild-type mice (Figures S6A and S6B). This finding suggests that LVe signaling influences the scaling of proprioceptive inputs to motor neurons.

DISCUSSION

The control of posture and balance is essential for motor performance. The vestibular system plays an important role in this pro-

cess through its ability to stabilize and adjust body position during movement (Angelaki and Cullen, 2008; Grillner et al., 1970; Wilson and Yoshida, 1968). Using genetic perturbation experiments, we demonstrate that signaling interactions between the proprioceptive and vestibular system play a key role in shaping connection specificity between vestibular neurons in the brainstem and spinal motor neurons. We discuss how these findings advance our understanding of vestibular system function, especially in the context of connectivity refinement and functional interaction with proprioceptive circuitry to ensure smooth motor behavior.

Motor Neuron Subtype Identity Aligns with Synaptic Input Specificity

Work on the cat lumbar spinal cord demonstrates that select lumbar extensor motor pools are favored direct synaptic targets for LVe input compared to flexor counterparts (Grillner et al., 1970), a profile we find to be conserved in mice. A key insight of our work is that the observed extensor-flexor interpool specificity pattern is supplemented by a preference of LVe input to target slow over fast motor neuron subtypes within each extensor pool analyzed, and notably, this bias is even detectable at the level of a general lumbar LMC motor neuron analysis.

What may be the functional reasons behind the identified vestibulo-motor connectivity profile to preferentially target slow over fast motor neurons within extensor pools? Vestibular input enhances the activation of motor neurons innervating extensor muscles exhibiting antigravitational function and can produce large motoneuronal depolarizations through temporal summation (Grillner et al., 1970). This is physiologically relevant since vestibular neurons fire at high frequencies (Angelaki and Cullen, 2008), also detected in awake behaving mice (Beranek and Cullen, 2007), demonstrating that the vestibular system has

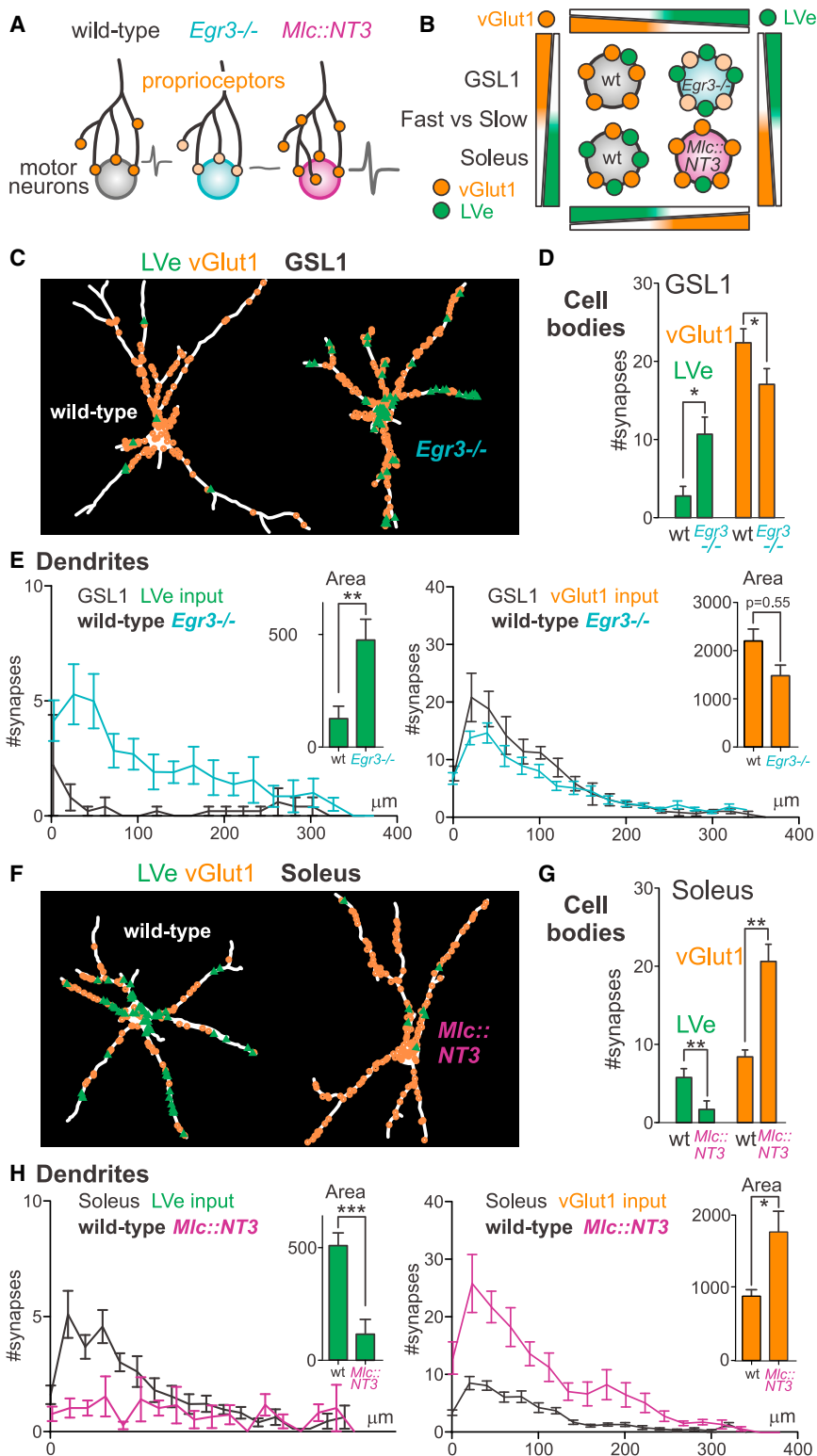


Figure 7. Muscle Spindle Signaling Influences Input to Motor Neuron Subtypes

(A) Synaptic input status of proprioceptors to alpha motor neurons in wild-type, *Egr3*^{-/-} and *MLC::NT3* mice. *Egr3*^{-/-} proprioceptive terminals are physically present but non-functional, whereas they show over-proliferation and aberrant connections in *MLC::NT3* mice.

(B) Summary diagram illustrating main findings. LVe inputs to motor neuron subtypes of extensor pools are affected by genetic manipulation of proprioceptor input function.

(C–H) NeuroLucida reconstruction (C and F) and quantification (D, E, G, and H) of LVe and vGlut1 synaptic input to GSL1 and soleus motor neurons in wild-type, compared to *Egr3*^{-/-} (for GSL1) and *MLC::NT3* (for soleus) mice both at the cell body (D and G) and dendrite (D and H) level. In (E) and (H) area under curves are quantified and shown in bar graphs.

See also Figures S4, S5, and S6.

subtypes selectively recruited during endurance and postural tasks and with the ability to support contractions without fatigue (Burke, 1967; Kanning et al., 2010). In contrast, fast motor neurons receive sparse direct vestibular input, in line with these motor neurons being recruited during fast and powerful muscle contractions but to fatigue quickly (Burke, 1967; Kanning et al., 2010).

Our work is focused on synaptic input specificity directly to motor neurons, but vestibular signaling also acts through indirect pathways via spinal interneurons, and these pathways also employ specific connectivity rules following motor pool-specific patterns (Grillner et al., 1970). Even though flexor motor pools do not receive direct excitatory LVe inputs, disynaptic pathways can specifically inhibit them and thereby further enhance the differential functional impact that LVe signaling exhibits on extensor and flexor motor pools. Electrophysiological studies on the organization of peripheral and rubrospinal inputs to motor neurons demonstrate that indirect inputs can also exhibit fiber-type-specific functional connectivity profiles (Burke et al., 1970). Whether indirect inputs to motor neurons in the vestibular system also

the capability to contribute to motoneuronal recruitment. Our work shows that the vestibular system contributes to this process by preferential targeting of slow extensor motor neuron

follow the intrapool motor unit twitch-type organizational principle as direct ones do will be an interesting question to pursue.

Multisensory Integration in the Motor Output System

The functionality of the motor system depends heavily on continuous integration of sensory information of different modalities. Multisensory inputs influence many neuronal elements along motor output pathways, a general organizational principle that is evolutionarily conserved even to circuits regulating *Drosophila* larvae behavior (Ohyama et al., 2015). Focusing on the last synapse of motor output pathways affecting movement, we found that vestibular and proprioceptive inputs converge on slow and fast motor neuron subtypes with an inverse anatomical synaptic scaling profile. While we favor the view that functional complementarity plays a role in synaptic scaling, whether similar scaling processes can also occur between functionally non-complementary inputs remains to be determined.

Our findings raise the question of how and where proprioceptive and vestibular systems interact functionally. Most relevant for our study, the vestibular system can enhance proprioceptive inputs in a synergistic manner (Grillner et al., 1970). Moreover, vestibular input to motor neurons inherently carries multisensory information. Vestibular neurons are secondary neurons in the chain of sensory input processing, receiving primary vestibular sensory input as well as indirect feedback from the proprioceptive and visual system (Angelaki and Cullen, 2008). In particular, somatosensory feedback circuits activated by passive hindlimb movement regulate vestibular neuron activity (Arshian et al., 2014). Thus direct vestibular input to motor neurons combines multiple sensory streams of different degrees of integration and we found that these inputs are organized into precise patterns and are complementary to direct proprioceptive inputs.

Developmental Mechanisms Guiding the Assembly of Inputs to Motor Neurons

The precise developmental assembly of synaptic inputs to motor neurons is a prerequisite for the functionality of the mature motor system. Despite its importance however, mechanistic insight exists for only a limited number of functionally defined neuronal subpopulations with synaptic access to motor neurons. The wiring specificity between proprioceptors and motor neuron pools within the same reflex arc is established early and through mechanisms independent of neuronal activity (Frank, 1990; Mears and Frank, 1997). Combinatorial action of neuronal and retrograde molecular factors as well as positional cues play important roles in instructing sensory-motor connectivity (Arber, 2012; Wenner and Frank, 1995). Yet sensory connectivity to synergistic motor pools refines at postnatal stages, a process influenced by proprioceptor neuron activity (Mendelsohn et al., 2015).

Here, we have assessed time course and mechanisms of vestibular input assembly and refinement to motor neurons. We found that while significant input differences between extensor (GS) and flexor (TA) motor neurons are already established at early postnatal stages, a likely activity-dependent postnatal synaptic refinement process abolishes vestibular input to TA motor neurons. We revealed that this process is driven at least in part by vestibular signaling itself. The time window during which refinement occurs (P11–P17) matches the emergence of posture and weight bearing in rodents (Geisler et al., 1993), raising the possibility that maturation of synaptic input may be linked to the emergence of postural behavioral abilities.

The second level of synaptic input scaling to motor neuron subtypes is shaped by bidirectional sensory signaling. Genetic manipulations affecting either the functionality of muscle spindle feedback or vestibular signaling resulted in adjustments of the other channel counterbalancing the genetically imposed changes. How could such input adjustment to motor neurons be regulated? We found that in mice, proprioceptive connections exhibit higher proximal synaptic input with gradually decreasing input on distal dendrites, in agreement with recent input reconstructions to rat motor neurons (Rotterman et al., 2014). Interestingly, compensatory LVe input distribution to GS1 FF motor neurons in *Egr3* mutant mice scales accordingly. Since these muscle spindle afferent synapses are present but non-functional (Chen et al., 2002), it is likely that the observed adjustment of synaptic input to motor neurons is not merely a competition for synaptic space. A plausible mechanism instead might be that synaptic input to motor neurons is regulated locally through retrograde and homeostatic mechanisms involving postsynaptic feedback from motor neurons. In this context, it is interesting to consider that individual group Ia afferents connect to almost all motor neurons supplying the same muscle (Mendell and Henneman, 1968). Ia input density scaling therefore likely occurs at the level of individual motor neurons according to subtype identity. Moreover, proprioceptor-driven vestibular synaptic scaling only operates on motor neuron pools to which LVe input has direct functional impact, as we observed no input scaling to TA motor neurons that receive proprioceptive but are devoid of LVe input. Thus, the two studied sensory channels differentially influence the refinement and scaling process of inputs to motor neurons, further supporting the idea that multiple independent layers regulate input specificity to motor neuron subtypes.

Bidirectional synaptic compensation may also explain at least part of the relatively minor locomotor phenotypes observed in *Egr3* mutants (Takeoka et al., 2014) and *R4^{Cre}::vGlut2^{fllox}* mice analyzed here, both of which exhibit signaling defects in the respective sensory system starting during development. Interestingly, cross-modal sensory regulation during development also appears to operate in humans. Patients with infant-onset vestibular system dysfunction show limited behavioral abnormalities likely due to somatosensory compensatory mechanism, whereas compensation following adult injury to the vestibular system is restricted (Horak et al., 1994). These observations suggest that there might be a developmentally defined critical period for cross-modal sensory regulation to adjust circuitry to motor neurons needed for posture and balance. Together, our work uncovers how sensory inputs of functionally complementary modality converge and influence each other at the final output step controlling movement, providing an important contribution to understanding specificity and function of the motor system.

EXPERIMENTAL PROCEDURES

Mouse Genetics

Tau^{lox-STOP-lox-Flp-INLA} (Pivetta et al., 2014), *Chodl^{PLAP}* (Sakurai et al., 2013), *NADPH oxidase 3 (Nox3)* mutant (Paffenholz et al., 2004), *vGlut2^{fllox}* (Jax Mice Strain #007583), *Egr3* mutant (Tourtellotte and Milbrandt, 1998), *R4^{Cre}* (Di Bonito et al., 2013), and *Mlc::NT3* (Wang et al., 2007) mouse strains were maintained on a mixed genetic background (129/C57Bl6). Housing, surgery,

behavioral experiments and euthanasia were performed in compliance with the Swiss Veterinary Law guidelines.

Virus Production and Injections

Rabies viruses (Rabies-mCherry and Rabies-GFP: Rabies-FP) used were amplified and purified from local viral stocks following established protocols (Stepien et al., 2010; Wickersham et al., 2007). All AAVs used in this study were described previously (Esposito et al., 2014; Pivet et al., 2014; Takeoka et al., 2014) and of genomic titers $>1 \times 10^{13}$. Additional information on anterograde and retrograde viral tracing, immunohistochemistry, imaging and anatomical quantification are found in the [Supplemental Experimental Procedures](#).

Behavioral Analysis

Open field, grip strength test, horizontal ladder locomotion, and beam tests were performed as previously described (Esposito et al., 2014; Takeoka et al., 2014; Carter et al., 2001). Additional information on detailed behavioral analyses is found in the [Supplemental Experimental Procedures](#).

Statistics

All statistical analysis, plots, and linear regression lines were made using GraphPad PRISM (v6.0). Column bar graphs and dot plots represent the average value \pm SEM. The means of different data distributions were compared using an unpaired Student's *t* test (Figures 1D, 1F, 1H, 2A, 2D, 3D, 3F, 4B, 4D, 5B–5D, 6A, 7D, 7G, S3, S4, S5B, and S6A). Correlation analysis was used for Figures 3E and 6B. A one-way ANOVA for independent measurements was used for comparing multiple TA datasets in Figure 4B. The area under the frequency-distribution curves in Figures 7I, 7H, S5, and S6B was used as a measure of synaptic input on dendrites. The correlogram plot shown in Figure 6C was obtained in R using the library ggplot. Significance level is defined as follows for all analyses performed: **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.023>.

ACKNOWLEDGMENTS

We are grateful to M. Mielich for expert technical help, M. Tripodi for help and guidance during initial stages of the project, F. Roselli for insights into motor neuron subtypes and muscle subcompartments, M. Studer for sharing *R4::Cre* mice with us, S. Bourke, C. Genoud, and L. Gelman from the FMI imaging facility, N. Ehrenfurchter from the Biozentrum Imaging facility and M. Stadler from the FMI Bioinformatics Platform for help and advice with image acquisition and analysis, and to P. Caroni and B. Roska for discussions and comments on the manuscript. E.B. was supported by a fellowship of the Werner Siemens Foundation, A.T. by an International Foundation for Research in Paraplegia (IRP) fellowship. All authors were supported by an ERC Advanced Grant, the Swiss National Science Foundation, the Kanton Basel-Stadt, and the Novartis Research Foundation.

Received: July 8, 2015

Revised: August 22, 2015

Accepted: September 1, 2015

Published: October 8, 2015

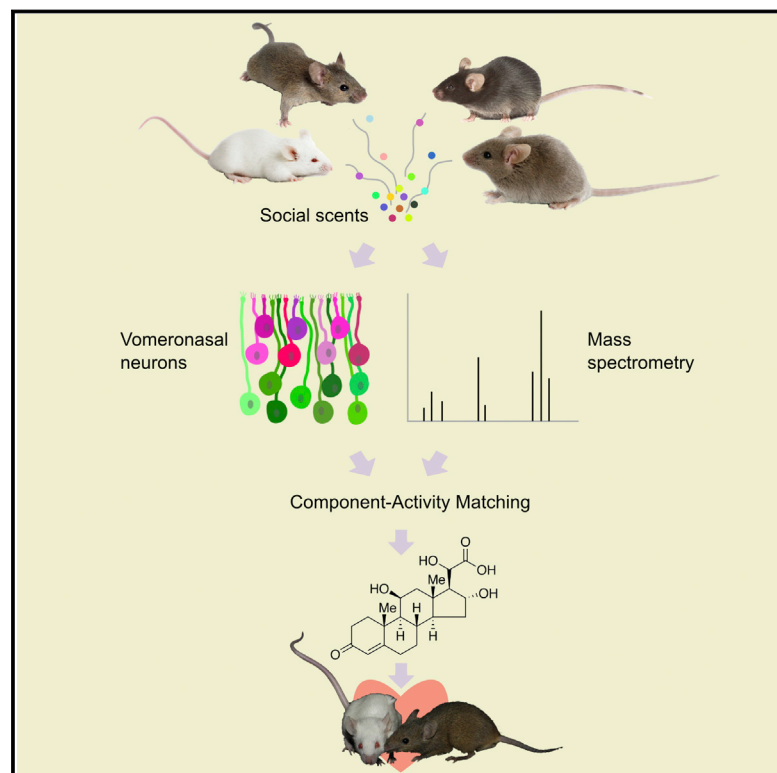
REFERENCES

- Abaira, V.E., and Ginty, D.D. (2013). The sensory neurons of touch. *Neuron* 79, 618–639.
- Angelaki, D.E., and Cullen, K.E. (2008). Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.* 31, 125–150.
- Arber, S. (2012). Motor circuits in action: specification, connectivity, and function. *Neuron* 74, 975–989.
- Arshian, M.S., Hobson, C.E., Catanzaro, M.F., Miller, D.J., Puterbaugh, S.R., Cotter, L.A., Yates, B.J., and McCall, A.A. (2014). Vestibular nucleus neurons respond to hindlimb movement in the decerebrate cat. *J. Neurophysiol.* 111, 2423–2432.
- Beraneck, M., and Cullen, K.E. (2007). Activity of vestibular nuclei neurons during vestibular and optokinetic stimulation in the alert mouse. *J. Neurophysiol.* 98, 1549–1565.
- Brodal, A., and Pompeiano, O. (1957). The vestibular nuclei in cat. *J. Anat.* 91, 438–454.
- Brown, A.G. (1981). *Organization of the Spinal Cord* (Springer Verlag).
- Burke, R.E. (1967). Motor unit types of cat triceps surae muscle. *J. Physiol.* 193, 141–160.
- Burke, R.E., Jankowska, E., and ten Bruggencate, G. (1970). A comparison of peripheral and rubrospinal synaptic input to slow and fast twitch motor units of triceps surae. *J. Physiol.* 207, 709–732.
- Burke, R.E., Dum, R.P., Fleshman, J.W., Glenn, L.L., Lev-Tov, A., O'Donovan, M.J., and Pinter, M.J. (1982). A HRP study of the relation between cell size and motor unit type in cat ankle extensor motoneurons. *J. Comp. Neurol.* 209, 17–28.
- Carter, R.J., Morton, J., and Dunnett, S.B. (2001). Motor coordination and balance in rodents. *Curr. Protoc. Neurosci. Chapter 8*, Unit 8.12.
- Chen, H.H., Tourtellotte, W.G., and Frank, E. (2002). Muscle spindle-derived neurotrophin 3 regulates synaptic connectivity between muscle sensory and motor neurons. *J. Neurosci.* 22, 3512–3519.
- Di Bonito, M., Narita, Y., Avallone, B., Sequino, L., Mancuso, M., Andolfi, G., Franzè, A.M., Puelles, L., Rijli, F.M., and Studer, M. (2013). Assembly of the auditory circuitry by a Hox genetic network in the mouse brainstem. *PLoS Genet.* 9, e1003249.
- Du Beau, A., Shakya Shrestha, S., Bannatyne, B.A., Jalicy, S.M., Linnen, S., and Maxwell, D.J. (2012). Neurotransmitter phenotypes of descending systems in the rat lumbar spinal cord. *Neuroscience* 227, 67–79.
- Eccles, J.C., Eccles, R.M., and Lundberg, A. (1957). The convergence of monosynaptic excitatory afferents on to many different species of alpha motoneurons. *J. Physiol.* 137, 22–50.
- Enjin, A., Rabe, N., Nakanishi, S.T., Vallstedt, A., Gezelius, H., Memic, F., Lind, M., Hjalt, T., Tourtellotte, W.G., Bruder, C., et al. (2010). Identification of novel spinal cholinergic genetic subtypes disclose Chodl and Pitx2 as markers for fast motor neurons and partition cells. *J. Comp. Neurol.* 518, 2284–2304.
- Esposito, M.S., Capelli, P., and Arber, S. (2014). Brainstem nucleus MdV mediates skilled forelimb motor tasks. *Nature* 508, 351–356.
- Frank, E. (1990). The formation of specific synaptic connections between muscle sensory and motor neurons in the absence of coordinated patterns of muscle activity. *J. Neurosci.* 10, 2250–2260.
- Geisler, H.C., and Gramsbergen, A. (1998). Motor development after vestibular deprivation in rats. *Neurosci. Biobehav. Rev.* 22, 565–569.
- Geisler, H.C., Westerga, J., and Gramsbergen, A. (1993). Development of posture in the rat. *Acta Neurobiol. Exp. (Warsz.)* 53, 517–523.
- Grillner, S., and Dubuc, R. (1988). Control of locomotion in vertebrates: spinal and supraspinal mechanisms. *Adv. Neurol.* 47, 425–453.
- Grillner, S., Hongo, T., and Lund, S. (1970). The vestibulospinal tract. Effects on alpha-motoneurons in the lumbosacral spinal cord in the cat. *Exp. Brain Res.* 10, 94–120.
- Horak, F.B., Shupert, C.L., Dietz, V., and Horstmann, G. (1994). Vestibular and somatosensory contributions to responses to head and body displacements in stance. *Exp. Brain Res.* 100, 93–106.
- Kanning, K.C., Kaplan, A., and Henderson, C.E. (2010). Motor neuron diversity in development and disease. *Annu. Rev. Neurosci.* 33, 409–440.
- Kaplan, A., Spiller, K.J., Towne, C., Kanning, K.C., Choe, G.T., Geber, A., Akay, T., Aebischer, P., and Henderson, C.E. (2014). Neuronal matrix metalloproteinase-9 is a determinant of selective neurodegeneration. *Neuron* 81, 333–348.

- Leroy, F., Lamotte d'Incamps, B., Imhoff-Manuel, R.D., and Zytnicki, D. (2014). Early intrinsic hyperexcitability does not contribute to motoneuron degeneration in amyotrophic lateral sclerosis. *eLife* 3, 3.
- Liang, H., Bácskai, T., Watson, C., and Paxinos, G. (2014). Projections from the lateral vestibular nucleus to the spinal cord in the mouse. *Brain Struct. Funct.* 219, 805–815.
- Liang, H., Bácskai, T., and Paxinos, G. (2015). Termination of vestibulospinal fibers arising from the spinal vestibular nucleus in the mouse spinal cord. *Neuroscience* 294, 206–214.
- Lund, S., and Pompeiano, O. (1968). Monosynaptic excitation of alpha motoneurons from supraspinal structures in the cat. *Acta Physiol. Scand.* 73, 1–21.
- Lundberg, A. (1975). Control of spinal mechanisms from the brain. In *The Basic Neurosciences* (Raven Press).
- Matthews, P.B. (1981). Evolving views on the internal operation and functional role of the muscle spindle. *J. Physiol.* 320, 1–30.
- Mears, S.C., and Frank, E. (1997). Formation of specific monosynaptic connections between muscle spindle afferents and motoneurons in the mouse. *J. Neurosci.* 17, 3128–3135.
- Mendell, L.M., and Henneman, E. (1968). Terminals of single Ia fibers: distribution within a pool of 300 homonymous motor neurons. *Science* 160, 96–98.
- Mendelsohn, A.I., Simon, C.M., Abbott, L.F., Mentis, G.Z., and Jessell, T.M. (2015). Activity regulates the incidence of heteronymous sensory-motor connections. *Neuron* 87, 111–123.
- Moorman, S.J., Cordova, R., and Davies, S.A. (2002). A critical period for functional vestibular development in zebrafish. *Dev. Dyn.* 223, 285–291.
- Ohyama, T., Schneider-Mizell, C.M., Fetter, R.D., Aleman, J.V., Franconville, R., Rivera-Alba, M., Mensh, B.D., Branson, K.M., Simpson, J.H., Truman, J.W., et al. (2015). A multilevel multimodal circuit enhances action selection in *Drosophila*. *Nature* 520, 633–639.
- Oliveira, A.L., Hydling, F., Olsson, E., Shi, T., Edwards, R.H., Fujiyama, F., Kaneko, T., Hökfelt, T., Cullheim, S., and Meister, B. (2003). Cellular localization of three vesicular glutamate transporter mRNAs and proteins in rat spinal cord and dorsal root ganglia. *Synapse* 50, 117–129.
- Paffenholz, R., Bergstrom, R.A., Pasutto, F., Wabnitz, P., Munroe, R.J., Jagla, W., Heinzmann, U., Marquardt, A., Bareiss, A., Laufs, J., et al. (2004). Vestibular defects in head-tilt mice result from mutations in *Nox3*, encoding an NADPH oxidase. *Genes Dev.* 18, 486–491.
- Pecho-Vrieseling, E., Sigrist, M., Yoshida, Y., Jessell, T.M., and Arber, S. (2009). Specificity of sensory-motor connections encoded by *Sema3e-Plxn1* recognition. *Nature* 459, 842–846.
- Pivetta, C., Esposito, M.S., Sigrist, M., and Arber, S. (2014). Motor-circuit communication matrix from spinal cord to brainstem neurons revealed by developmental origin. *Cell* 156, 537–548.
- Pun, S., Santos, A.F., Saxena, S., Xu, L., and Caroni, P. (2006). Selective vulnerability and pruning of phasic motoneuron axons in motoneuron disease alleviated by CNTF. *Nat. Neurosci.* 9, 408–419.
- Rotterman, T.M., Nardelli, P., Cope, T.C., and Alvarez, F.J. (2014). Normal distribution of VGLUT1 synapses on spinal motoneuron dendrites and their reorganization after nerve injury. *J. Neurosci.* 34, 3475–3492.
- Sakurai, K., Akiyama, M., Cai, B., Scott, A., Han, B.X., Takatoh, J., Sigrist, M., Arber, S., and Wang, F. (2013). The organization of submodality-specific touch afferent inputs in the vibrissa column. *Cell Rep.* 5, 87–98.
- Shinoda, Y., Ohgaki, T., Futami, T., and Sugiuchi, Y. (1988). Vestibular projections to the spinal cord: the morphology of single vestibulospinal axons. *Prog. Brain Res.* 76, 17–27.
- Stepien, A.E., Tripodi, M., and Arber, S. (2010). Monosynaptic rabies virus reveals premotor network organization and synaptic specificity of cholinergic partition cells. *Neuron* 68, 456–472.
- Takeoka, A., Vollenweider, I., Courtine, G., and Arber, S. (2014). Muscle spindle feedback directs locomotor recovery and circuit reorganization after spinal cord injury. *Cell* 159, 1626–1639.
- Tourtellotte, W.G., and Milbrandt, J. (1998). Sensory ataxia and muscle spindle agenesis in mice lacking the transcription factor *Egr3*. *Nat. Genet.* 20, 87–91.
- Uchino, Y., and Kushiro, K. (2011). Differences between otolith- and semicircular canal-activated neural circuitry in the vestibular system. *Neurosci. Res.* 71, 315–327.
- Van Cleave, S., and Shall, M.S. (2006). A critical period for the impact of vestibular sensation on ferret motor development. *J. Vestib. Res.* 16, 179–186.
- Walton, K.D., Harding, S., Anschel, D., Harris, Y.T., and Llinás, R. (2005). The effects of microgravity on the development of surface righting in rats. *J. Physiol.* 565, 593–608.
- Wang, Z., Li, L.Y., Taylor, M.D., Wright, D.E., and Frank, E. (2007). Prenatal exposure to elevated NT3 disrupts synaptic selectivity in the spinal cord. *J. Neurosci.* 27, 3686–3694.
- Wenner, P., and Frank, E. (1995). Peripheral target specification of synaptic connectivity of muscle spindle sensory neurons with spinal motoneurons. *J. Neurosci.* 15, 8191–8198.
- Wickersham, I.R., Lyon, D.C., Barnard, R.J., Mori, T., Finke, S., Conzelmann, K.K., Young, J.A., and Callaway, E.M. (2007). Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* 53, 639–647.
- Wilson, V.J., and Yoshida, M. (1968). Vestibulospinal and reticulospinal effects on hindlimb, forelimb, and neck alpha motoneurons of the cat. *Proc. Natl. Acad. Sci. USA* 60, 836–840.
- Windhorst, U. (2007). Muscle proprioceptive feedback and spinal networks. *Brain Res. Bull.* 73, 155–202.

A Molecular Code for Identity in the Vomeronasal System

Graphical Abstract



Authors

Xiaoyan Fu, Yuetian Yan, Pei S. Xu, ..., Wongi Chong, Michael L. Gross, Timothy E. Holy

Correspondence

holy@wustl.edu

In Brief

Component-activity matching (CAM), a technique for selecting candidate ligands that “explain” patterns of bioactivity across diverse complex mixtures, enabled the discovery of several new ligands that explain firing rates in functionally distinct classes of vomeronasal neurons, including a new set of ligands that are female sex pheromones for the mouse.

Highlights

- CAM is an approach to match candidate ligands and specific patterns of bioactivity
- 23 compounds in mouse urine explain firing rates in 7 classes of vomeronasal neurons
- Steroid carboxylic acids are a new family of vomeronasal ligands
- Steroid carboxylic acids are female sex pheromones for the mouse

A Molecular Code for Identity in the Vomeronasal System

Xiaoyan Fu,¹ Yuetian Yan,² Pei S. Xu,¹ Ilan Geerloff-Vidavsky,² Wongi Chong,¹ Michael L. Gross,² and Timothy E. Holy^{1,*}

¹Department of Anatomy and Neurobiology, Washington University in St. Louis, St. Louis, MO 63110, USA

²Department of Chemistry, Washington University in St. Louis, St. Louis, MO 63110, USA

*Correspondence: holy@wustl.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.012>

SUMMARY

In social interactions among mammals, individuals are recognized by olfactory cues, but identifying the key signals among thousands of compounds remains a major challenge. To address this need, we developed a new technique, component-activity matching (CAM), to select candidate ligands that “explain” patterns of bioactivity across diverse complex mixtures. Using mouse urine from eight different sexes and strains, we identified 23 components to explain firing rates in seven of eight functional classes of vomeronasal sensory neurons. Focusing on a class of neurons selective for females, we identified a novel family of vomeronasal ligands, steroid carboxylic acids. These ligands accounted for much of the neuronal activity of urine from some female strains, were necessary for normal levels of male investigatory behavior of female scents, and were sufficient to trigger mounting behavior. CAM represents the first step toward an exhaustive characterization of the molecular cues for natural behavior in a mammalian olfactory system.

INTRODUCTION

Mammals explore the chemical world with several olfactory modalities (Ma, 2007). The accessory olfactory system (AOS) emphasizes the detection of social cues, sometimes called pheromones, that regulate behavior among members of the same species (Dulac and Torello, 2003; Halpern and Martínez-Marcos, 2003). One of the great attractions of the AOS is the opportunity to explore the molecular, cellular, and circuit underpinnings of behavior in a genetically tractable mammal. Perhaps the most substantial current barrier to exploit this promise is our incomplete understanding of the natural cues that the AOS detects. Knowing more about the nature of the stimuli will enable sophisticated studies of sensory coding, circuit function, and behavior.

Several scent sources are known to be involved in chemical communication in mammals: secretory glands (lacrimal, salivary, and preputial), urine, and feces (Halpern and Martínez-Marcos, 2003; Kimoto et al., 2005). Mouse urine excites widespread activity among vomeronasal sensory neurons (VSNs) (He et al.,

2008; Holy et al., 2000; Stowers et al., 2002; Tolokh et al., 2013) and is the best behaviorally characterized source of chemical cues for mammalian social communication (Halpern and Martínez-Marcos, 2003). Mouse urine conveys information about the sex and strain (Brennan and Keverne, 1997; Kimchi et al., 2007; Leypold et al., 2002; Pankevich et al., 2004; Stowers et al., 2002). Although progress has been made recently toward identifying the molecular nature of pheromone cues by purifying individual ligands from mouse urine (Chamero et al., 2007; Hsu et al., 2008; Nodari et al., 2008), the identities of olfactory ligands inside urine cues are largely unknown; there is not even an estimate of how many distinct compounds comprise the “olfactory identity” of an individual.

Here, we developed an approach, component-activity matching (CAM) (Figure 1), with the aim of systematically and exhaustively defining a “short list” of candidate vomeronasal ligands that encode identity. This approach exploits the striking differences among natural stimuli across different sexes and strains of mice (Tolokh et al., 2013), which we use as a proxy for the variability that one might find at the individual level in natural, genetically diverse, populations of mice. With CAM, we performed a forward screen for compounds that may drive activity in vomeronasal neurons, without the need for laborious sample purification. Using a combination of physiological recording and quantitative liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Figure 1), we identified a small set of constituents whose concentrations match a pattern of neuronal responsiveness across samples.

Using this approach, we focused on a class of neurons that responded to urine from females of all strains, but not to any male strains. We purified and structurally identified the two best CAM candidates to explain this pattern of neuronal activity. These two candidates are carboxylic acid steroid metabolites, which we call “cortigynic acid” and “corticosteronic acid.” Cortigynic acid accounts for one-fourth of neuronal activity of female C57BL/6J mouse urine. Cortigynic acid was 7-fold more abundant in the urine of gonadally intact adult female mice than in juvenile mice, was 30-fold higher than in an ovariectomized female, and was not detectable in male mice. When added to the urine of male mice, cortigynic acid induced additional exploratory behavior as observed previously with whole female urine (Pankevich et al., 2004; Holy and Guo, 2005; Guo and Holy, 2007). Removing cortigynic acid from the female urine extract significantly decreased male’s investigation time toward the “deficient” female urine extract. Finally, we found that the combination of cortigynic

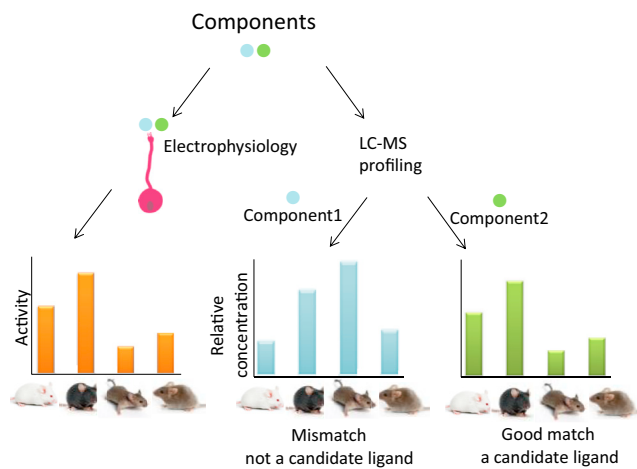


Figure 1. CAM for Ligand Identification

A diverse collection of samples (here, urine extracts from male and female mice of different strains) are analyzed to extract the relative concentrations of each component by LC-ESI mass spectrometry and subjected to an assay of activity (left, by recordings of spiking responses). For a selected neuron, physiological responses (neuronal firing rate, orange bars) to the different stimuli are compared to the abundances of individual components (blue and green bars). Component 1 (blue bars) is not distributed in a manner that could explain the firing rate responses and is therefore an implausible candidate. In contrast, component 2 (green bars) has an across-sample distribution consistent with the measured neuronal firing rates and is thus a plausible candidate to explain the response.

acid and corticosterone is sufficient to promote male-mourning behavior.

This manuscript describes the first steps in a comprehensive analysis of the molecular cues for identity in a mammalian olfactory system. We applied this approach to a long-standing mystery (Dixon and Mackintosh, 1975) and discovered the chemical identity of female sex pheromones for the mouse.

RESULTS

A New Strategy for Identifying Vomeronasal Ligands from Complex Mixtures

In urine, much of the activity that drives firing in VSNs originates from multiple compounds with similar physico-chemical properties (Nodari et al., 2008). This similarity poses a major challenge to traditional methods for identifying the source of neuronal activity from complex mixtures. To overcome the difficulties, we developed a new strategy based on the observation that urine from different sexes and strains excites differing patterns of activity in vomeronasal neurons (Tolokh et al., 2013). We hypothesized that these different neuronal patterns could be caused by different ligands present in one or more of urine samples. We reasoned that the molecular correlates of identity might, therefore, be discoverable by seeking compounds whose abundance could “explain” these patterns of physiological responses. By picking out compounds whose concentrations across samples matched the pattern of neuronal firing (activity), we were able to identify plausible VSN candidate ligands without any preliminary purification (Figure 1).

Diverse Vomeronasal Responses to Urine Samples from Different Sexes and Strains

Inspired by the genetic differences among individuals present in natural populations, we selected mice of four laboratory strains (BALB/c, C57BL/6J, CBA, and DBA; Figure 1) and tested their C18 urine extracts for their ability to drive spiking responses in mouse vomeronasal neurons. We used a multi-electrode array (Nodari et al., 2008; Amson et al., 2010; Holy et al., 2000) to record spiking responses of large numbers of VSNs simultaneously from 13 intact vomeronasal neuroepithelia isolated from B6D2F1 male mice (Figure 2A). To reduce potential confounds from saturation, each of these eight urine extracts was tested at three different dilutions: 1:100, 1:300, and 1:1,000. A total of 488 single neurons were isolated from these spike recordings. To test the reproducibility of the responses, each stimulus was presented once in each of five cycles, each cycle delivering stimuli and negative controls alternating with flush solution in pseudo-random order. This stimulation paradigm resulted in a very large number—a total of 130—stimulus trials for each neuron in each preparation (Figure 2A). Responses were quantified in terms of the mean change in firing rate, Δr , upon stimulation (Figure 2B).

Of 488 single units, 121 responded ($\Delta r > 5$ Hz and distinguishable from the response to the standard at a level of $p < 0.05$, t test) to at least one stimulus (Figure S1). No neuron passing this test responded significantly to the independent negative control (Figure S1). Individual neurons showed a diversity of responses: some showed specificity to a single sample; others responded to a few strains from the same sex; and yet others responded selectively for one or the other sex but were relatively insensitive to strain. Examination of the entire population of recorded neurons revealed that functionally similar neurons were observed repeatedly across different preparations (Figure 2C). Of 121 responsive single units, we manually collected 57 of the cells into eight functional classes, based on specificity, sensitivity, and repeated observation of patterns across multiple single units. Of the remaining (not categorized) units, most were weakly responsive, and the rest exhibited idiosyncratic responses that did not, on their own, appear to clearly justify formation of a class (Figure S1). These classes served to select particularly clear “exemplar” neurons, which we temporarily assume to represent the entire class, for the purpose of identifying the ligands that drive them.

Chemical Profiles across Sex and Strain

These neurophysiological responses are presumably attributable to the presence of particular ligands in one or more of these samples, each at a specific concentration. If a response is driven by a single ligand, its pattern should be predictable from the concentration of the ligand across samples (Figure 1)—samples containing higher concentrations of this ligand should drive more activity. We reasoned that the molecular correlates of identity might, therefore, be discoverable by seeking compounds whose abundance could “explain” these patterns of physiological response.

To quantify the relative abundance of the individual molecular components, these same urine extracts were profiled with LC-MS/MS. To achieve high sensitivity and specificity, we performed negative-ion nanoLC-MS/MS using a water/acetonitrile

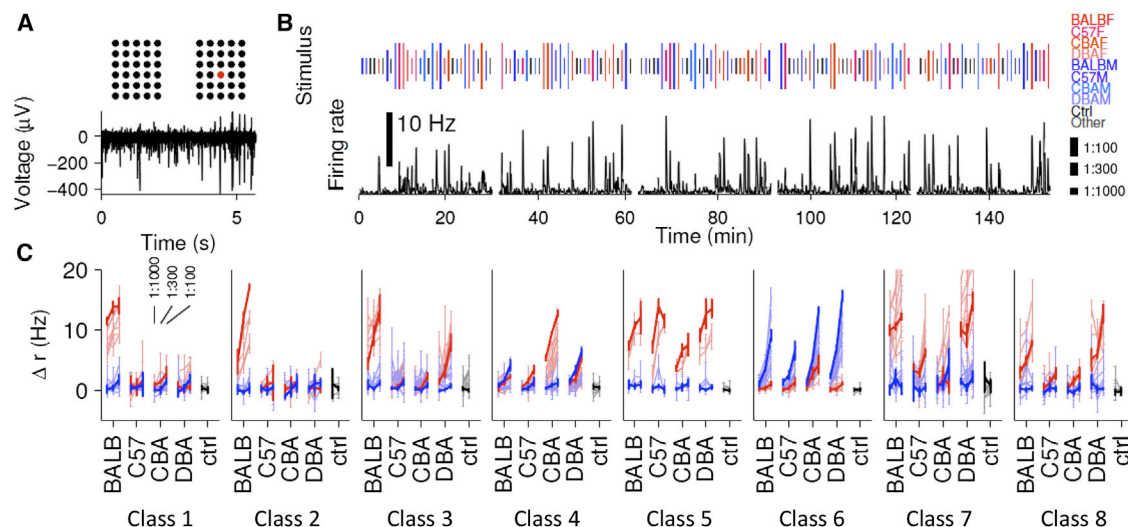


Figure 2. Categorization of VSN Responses

(A) Electrophysiological recording with a multi-electrode array. The voltage signal recorded from a representative electrode (indicated by red color) is illustrated. (B) Experimental design: firing rate versus time of one neuron, across five cycles of stimulus presentation. A collection of extracts from males (M) and females (F) of four strains, encoded by color, are presented at three different concentrations (encoded by stimulus bar height). Stimulus order is randomized across cycles, and each stimulus is presented once per cycle. Note that collecting this number of sensory responses required more than 2.5 hr of continuous recording.

(C) Overlays of all neurons assigned to one of eight repeatedly identified response patterns. Single exemplar neurons are shown in bold colors, and other neurons of similar functional type are shown in faint colors. Red indicates responses to female urine; blue indicates responses to male; and responses to all three dilutions are shown. The primary differentiator of the first two classes is their concentration dependence. Error bars represent firing rate SEM.

See also Figure S1.

gradient for the LC and high mass resolving power (to enable mass accuracy of a few ppm). It is apparent (Figure S2A) that these samples show high chemical diversity, consistent with the differences observed electrophysiologically. When viewed at high mass resolving power, individual constituents were readily isolated from one another (Figures S2A and S2B). In some cases, multiple compounds with the same mass-to-charge (m/z) ratio eluted at different times. We used a combination of automated and manual algorithms to segment individual peaks (see Experimental Procedures), focusing on those representing components with reasonably high ion counts in at least one sample. Altogether, we quantified the integrated ion counts for components represented by 1,634 different peaks with 862 distinct m/z values. In terms of total ion count, these components spanned more than four orders of magnitude in abundance (Figure S2C), thus representing a “deep” profile of the individual constituents of these samples.

Molecular Correlates of Neuronal Responses

The central hypothesis of CAM is that individual compounds, present in one or more of the urine samples, might be the source of individual VSN responses. To test this hypothesis, we combined the relative abundance of a single peak across samples with the dilution factor during stimulus delivery, thereby generating an “effective concentration” of this putative ligand across all samples tested electrophysiologically. We then examined the extent to which the pattern of concentrations across samples could explain a VSN’s response (Figure 1). We fit the responses to the Hill equation (Hill, 1938; Aronson and Holy, 2013), quanti-

fying the fitting error by χ^2 (see Experimental Procedures). Thus, each VSN/component pair yielded a single χ^2 measuring the degree to which a particular ligand might “explain” this neuron’s responses. Collectively, taking one exemplar from each of the eight neuronal classes (Figure 2C) led to 13,072 distinct measures of goodness of fit.

An example class 2 (Figure 2C, marked with yellow arrow in Figure S1) neuron (shown in Figure 3) exhibited a higher firing rate to BALB/c females, even at 1:1,000 dilution, than to any other sample (Figure 3A). This implies that any ligand explaining this pattern of firing must have at least 10-fold higher concentration in this sample than in any other sample. Most compounds did not fit this pattern. To illustrate, we choose one example, a compound giving an ion of $m/z = 336.0723$ eluting at 34%–37% acetonitrile (Figure 3B) and exhibiting relatively little selectivity across samples. Correspondingly, when the concentration of this component, incorporating the dilution factor, was paired with firing rate, the resulting pattern of firing across samples was a poor fit to this model (Figure 3C, $\chi^2 = 352.6$) and could be confidently excluded ($p < 10^{-50}$, 21 degrees of freedom). In contrast, this constraint was met for a component giving an ion of $m/z = 427.1793$ eluting at 52%–56% acetonitrile (Figure 3D), and the firing pattern was well fit by the Hill equation with $\chi^2 = 26.6$ (Figure 3E; $p = 0.18$, meaning that the fit cannot be excluded as an explanation of this pattern of firing).

We performed this analysis for all eight “exemplar” neurons (one per class), each with all 1,634 components. For the exemplar class 1 neuron (Figure 2C, marked with green arrow in Figure S1), the fitting error to all components is shown in Figure 3F.

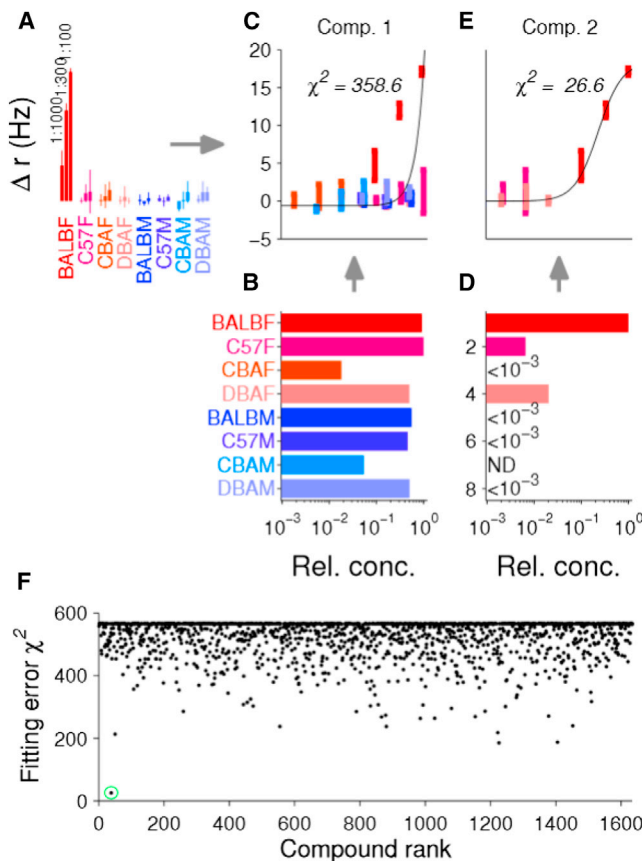


Figure 3. Matching Component Concentrations to Single-Neuron Activity Patterns

(A) Firing rate of a neuron from the second class in Figure 2C, marked with a yellow arrow in Figure S1.
 (B) Relative abundance across samples of a component with m/z 336.0723 eluting between 34%–37% acetonitrile.
 (C) Incorporation of the abundance with the dilution factor for each stimulus leads to a putative dose-response curve as a function of the concentration of this particular component. Fitting the points to a Hill model leads to a goodness-of-fit measured by χ^2 . This example represents a very poor fit ($p < 10^{-50}$).
 (D) Another component with m/z 427.1793 eluting between 52%–56% acetonitrile is most abundant in BALB/c female urine.
 (E) The dose-response curve from this second component fits well ($p = 0.18$).
 (F) χ^2 for a class 1 neuron (Figure 2C, marked with a green arrow in Figure S1) for each of 1,634 components. Only a single component (marked with a green circle), the one in (D) and (E), represents a good fit ($\chi^2 < 75$). See also Figure S2.

Remarkably, only one gave a plausible fit, namely, the compound giving an ion of $m/z = 427.1793$ shown in Figures 3D and 3E. This accurate mass is consistent with a compound, 4-pregnene-11,20,21-triol-3-one 21-sulfate (Figure 3F, marked with a green circle), that we previously identified from BALB/c female mouse urine (Hsu et al., 2008; Nodari et al., 2008).

This result is a striking demonstration of CAM's ability to identify plausible candidate ligands among the thousands of compounds in mouse urine. To analyze all eight classes of physiological responses (Figure 2C), we set a generous cutoff of $\chi^2 < 75$ and focused on components represented by 400 peaks with

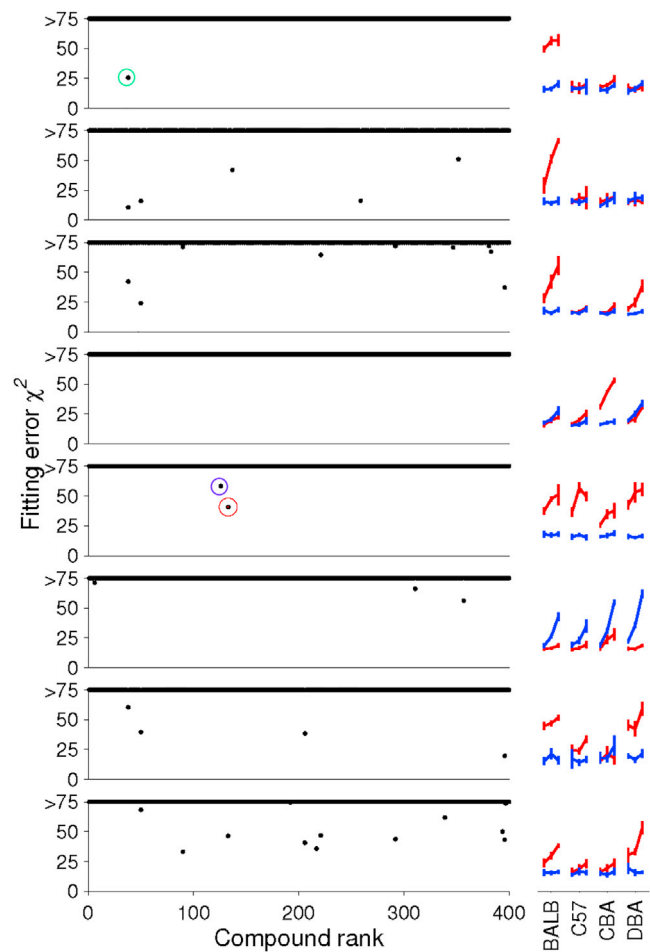


Figure 4. Candidate Ligands for Each Class of VSN Response

Goodness-of-fit plotted for the 400 most abundant components for each neuron shown at right. In each case, at most a handful of components represent plausible ligands. For each candidate with $\chi^2 < 75$, m/z and elution data are listed in Table S1. Circles highlight a previously identified CAM ligand m/z 427.1793 (4-pregnene-11,20,21-triol-3-one 21-sulfate, green) and the two novel CAM candidates to explain a pan-female response (class 5), m/z 377.1971 (red) and m/z 361.2018 (purple). See also Figure S3 and Table S1.

highest ion count, as these were most likely to be reliably quantified across samples and purified in sufficient amounts for further experiments. For 7/8 sensory neuron classes, we obtained one or more plausible candidate ligands (Figure 4; with 1, 5, 9, 0, 2, 3, 4, and 12 candidates for classes 1–8, respectively). Altogether, this approach identified 23 unique candidate ligands to explain these seven patterns of response.

The entire list of candidates is presented in Table S1. These 23 components of urine represent a first short list of molecular cues encoding an individual animal's "vomeronasal identity."

Isolation of CAM Candidate Ligands for "Pan-Female" Neuron Response Type

To test whether CAM successfully identified novel ligands, we decided to focus on a class of neuronal response lacking any

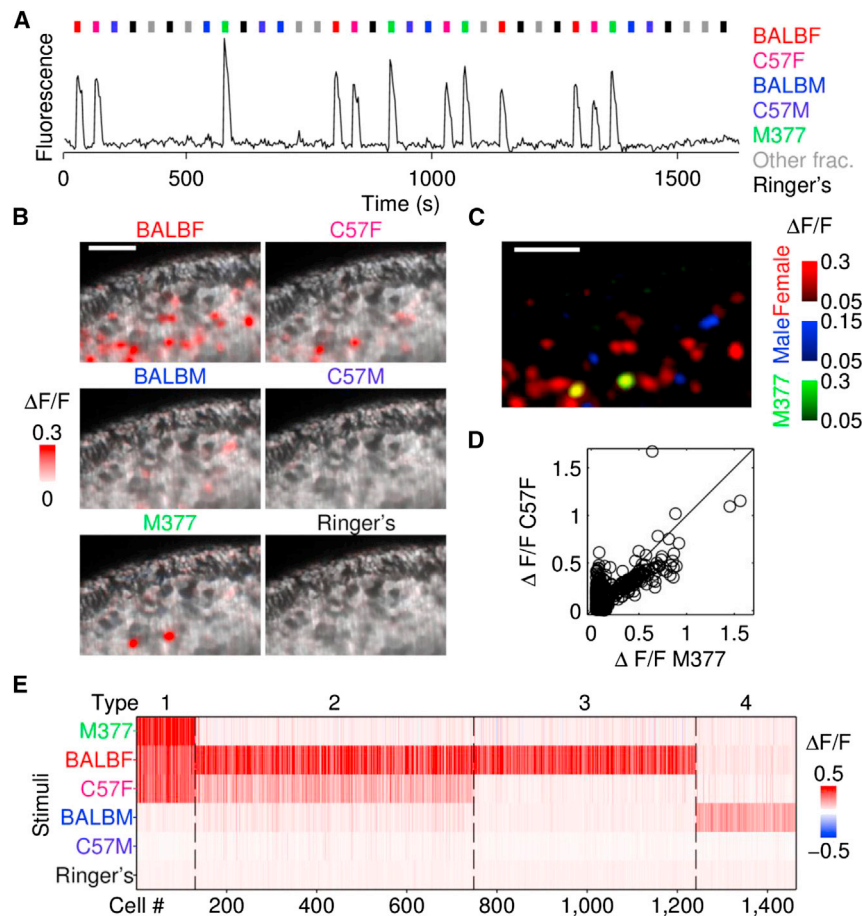


Figure 5. M377 Excites Female-Selective Neurons

(A) Experimental paradigm for calcium-imaging experiments. Each stimulus presentation is marked by a color-coded bar, each repeated four times. GCaMP2 fluorescence intensity for a particular neuron is shown.

(B) Single optical sections of the whole-mount VNO, with the dendritic knob layer near the top of the image. $\Delta F/F$ is encoded voxel-wise using a red color scale. Female extracts excite more neuronal activity than male extracts; M377 potentially excites a subset of neurons. Scale bar, 50 μm .

(C) Overlay of $\Delta F/F$ responses across stimuli. Note responses to M377 overlap responses to those cells activated by female cues.

(D) Comparison of response amplitudes to C57BL/6J female extract and purified M377. Each neuron sensitive to C57BL/6J female extract is shown as a separate circle. Diagonal line represents equal response to the two stimuli.

(E) Heatmap of all urine- or M377-responsive neurons imaged in two male VNOS. Each cell is represented as a single column, and cells with similar response patterns were grouped. The first three groups correspond to three different types of female-selective responses. Note that all neurons showing roughly equal responsiveness to both BALB/c and C57BL/6J female urine are also responsive to M377. See also Figure S4.

known ligands. Among mice, pheromones and other social odor cues convey information about sex, social status, and identity. We selected the “pan-female” response type (the fifth class in Figure 2C), as no VSN ligand for “femaleness” has yet been identified. This pattern of neuronal activity had just two likely candidates, of which the best match was a compound that we preliminarily denoted as M377; the component elutes at 44%–46% acetonitrile and gives by mass spectrometry an anion of m/z 377.1971 (presumably $[M - H]^-$, marked with a red circle in Figure 4) in the negative-ion mode. Direct inspection of the nanoLC-MS/MS profiles confirmed that this compound was found specifically in female mouse urine (Figure S3). This class of neurons has just one other CAM candidate shown in Figure 4, a compound we called M361, eluting at 54%–58% acetonitrile with a m/z 361.2018 (marked with a purple circle in Figure 4).

The LC data indicated that these two compounds co-elute with many others, but that these contaminants are less severe in samples from C57BL/6J females. Therefore, we chose this strain for further purification. We fractionated 100 ml of C57BL/6J female mouse urine extract by high-performance liquid chromatography (HPLC) (see Experimental Procedures), using mass spectrometry to identify fractions containing ions of the anticipated m/z 377 and 361. The corresponding HPLC fractions were further purified by thin-layer chromatography (TLC) to obtain a sample of high purity.

M377 Accounts for Approximately One-Fourth of the Neuronal Response to Urine from C57BL/6J Females

To determine whether M377 is the source of activity for all “class 5” neurons (Figure 4) as predicted by CAM, we used OCPI/light-sheet fluorescence microscopy (Holekamp et al., 2008) and calcium imaging to monitor neuronal activity in VSNs. Using male mice expressing GCaMP2 in VSNs (He et al., 2008), we recorded responses in the intact whole-mount preparation (Turaga and Holy, 2012; Xu and Holy, 2013). We presented urine extracts from BALB/c and C57BL/6J males and females, as well as purified M377 matching its endogenous concentration in C57BL/6J female urine extract. These stimuli were presented in four repeated trials, randomizing the order of stimuli within each cycle (Figure 5A). We recorded from two epithelial imaging volumes, containing an estimated 20,000 vomeronasal neurons (Turaga and Holy, 2012).

We identified 1,463 cells as responding robustly to at least one of the urine extracts. The largest numbers of neurons were activated by urine from females, consistent with electrophysiological recordings (Figure S1). Individual neurons were largely sex specific and displayed a range of strain selectivity. We also observed 130 neurons responsive to purified M377, demonstrating that this compound indeed activates a subset of neurons in the vomeronasal organ (VNO). Moreover, neurons responsive to M377 were also invariably responsive to urine from C57BL/6J and

BALB/c females, but did not respond to urine from males (Figures 5B, 5C, and 5E).

To investigate quantitatively the relationship between responses to M377 and mouse urine, for each M377-responsive neuron, we compared the intensity of its $\Delta F/F$ to both M377 and C57BL/6J female mouse urine. The response amplitudes were closely matched over the entire range of response intensities (Figure 5D), indicating a 1:1 relationship between these neurons' responses to C57BL/6J female mouse urine and purified M377. Consequently, we conclude that these neurons' urine responses are caused by M377.

The characteristics of the M377-responsive neurons shown in Figure 5E—approximate equal sensitivity to BALB/c and C57BL/6J female urine—are consistent with class 5 neurons (Figures 2C and 4). Two other types of female-responsive neurons, labeled types 2 and 3 (Figure 5E), are consistent with classes 7–8 and 1–3 (Figure 4), respectively; these neurons were unresponsive to M377.

Structural Characterization of M377

Based on accurate mass measurement, M377 is ionized to an $[M - H]^-$ chemical formula of $C_{21}H_{29}O_6^-$, indicating that it is a novel ligand. Following determination of its chemical formula, we solved its structure by using a combination of MS-based methods, including high mass resolving power, tandem MS (MS^2 and MS^3), HD exchange (to count exchangeable hydrogens), and a chemical reaction with periodic acid, followed by MS detection. The details of structure elucidation are provided in Supplemental Experimental Procedures and Figures S4, S5, S6 and Table S2. As a result of these analyses, the proposed structure of M377 is 16-hydroxycorticosterone 20-hydroxy-21-acid.

To test the validity of this structural interpretation, we commissioned a custom synthesis of this proposed compound, as an epimeric mixture at carbon 20. Urine-derived M377 and synthetic 16-hydroxycorticosterone-20-hydroxy-21-acid were indistinguishable by thin-layer chromatography (TLC) and reversed-phase HPLC (Figure S4A). Further, product-ion mass spectra of the synthetic and endogenous compounds were nearly indistinguishable (Figures 6A and 6B), indicating identical structures of the two. Fragmentations by MS^3 of major MS^2 fragment ions from both compounds are also consistent (data not shown).

We then asked whether synthetic M377 (16-hydroxycorticosterone 20-hydroxy-21-acid) showed the same neuronal activity as endogenous M377. The two stereoisomeric components of the synthetic epimeric mixture were quantified by HPLC as having a concentration ratio of 10:1, of which the less-abundant component matched the endogenous compound in elution time (Figure S4A). Synthetic 16-hydroxycorticosterone-20-hydroxy-21-acid and endogenous M377 were presented at three different concentrations (1, 0.1, and 0.01 μM) in four repeated trials, randomizing the order of stimuli within each cycle. VSNs responding to the synthetic compound also responded to endogenous M377 with identical concentration dependence (Figure 6C). These results conclusively demonstrate that M377 is 16-hydroxycorticosterone-20-hydroxy-21-acid.

This structure has several distinctive features. To our knowledge, biological activity of a steroid carboxylic acid has not been previously reported. As a 21-carbon steroid with a hydroxyl

group at the 11th position, this structure likely derives from metabolism of a glucocorticoid (Wang et al., 2010). However, the hydroxyl group at position 16 is a motif characteristic of a female sex steroid, estriol, which in humans is only produced in significant amounts during pregnancy (Raju et al., 1990). We therefore named this new VSN ligand “cortigynic acid.”

Structural Characterization of M361

Based on accurate mass measurement, M361 gives an $[M - H]^-$ chemical formula of $C_{21}H_{29}O_5^-$, indicating that it is a novel ligand. By comparing its chemical formula with that of cortigynic acid ($C_{21}H_{29}O_6^-$), we deduced that this new unknown compound might be an immediate metabolic precursor of cortigynic acid. We hypothesized that the distinctive C-16 hydroxyl group of M377 might be added as the last step, and therefore, we proposed that M361 is 20-dihydro-corticosterone-21-carboxylic acid. To test our proposed structure for M361, we also commissioned a custom synthesis of this second compound. Urine-derived M361 and synthetic 20-dihydro-corticosterone-21-carboxylic acid were indistinguishable by thin-layer chromatography (TLC) (data not shown) and reversed-phase HPLC (Figure S4B). MS/MS analysis demonstrated that the product-ion mass spectra of the synthetic and endogenous compounds are identical (Figures 6E and 6F). These results suggest that M361 is 20-dihydro-corticosterone-21-carboxylic acid, which we will also refer to as corticosteronic acid.

Combinatorial Coding of Female-Specific Ligands

To learn more about how these ligands are encoded by VSNs, we examined responses to a stimulus battery that included cortigynic acid and corticosteronic acid, together with 100-fold diluted urine samples from all four strains (BALB/c, DBA, C57BL/6J, and CBA) of females. Cortigynic acid and corticosteronic acid were presented individually at their endogenous concentrations in whole and in undiluted urine and in mixtures of the two spanning two orders of magnitude in concentration, the lowest of these approximately matching their concentrations in the diluted urine extract samples. Using male mice expressing GCaMP3 in VSNs, we recorded responses in the intact whole-mount preparation (Turaga and Holy, 2012; Xu and Holy, 2013) by OCPI microscopy (Holekamp et al., 2008).

From three epithelial imaging volumes, we identified 236 cells responding robustly to either cortigynic acid or corticosteronic acid. Using a clustering algorithm, we identified four types of responses. The first type was highly responsive to all four strains of females (the signature of class 5 neurons; Figures 2C and 4) and exquisitely sensitive to both cortigynic acid (M377) and corticosteronic acid (M361), exhibiting responses at sub-micromolar concentrations as present in the diluted urine samples. With the exhaustive sampling of OCPI microscopy and expanded range of stimulus concentrations, we also identified three other clusters responsive to at least one of these compounds and a subset of the mouse strains. One of these was strongly activated by BALB/c, DBA, and C57BL/6, and perhaps weakly by CBA females; these neurons were responsive to high concentrations of M377 and M361, but only weakly responsive at the lowest concentration. As CBA females have the lowest endogenous concentration of M377 (by a factor of approximately 2; see

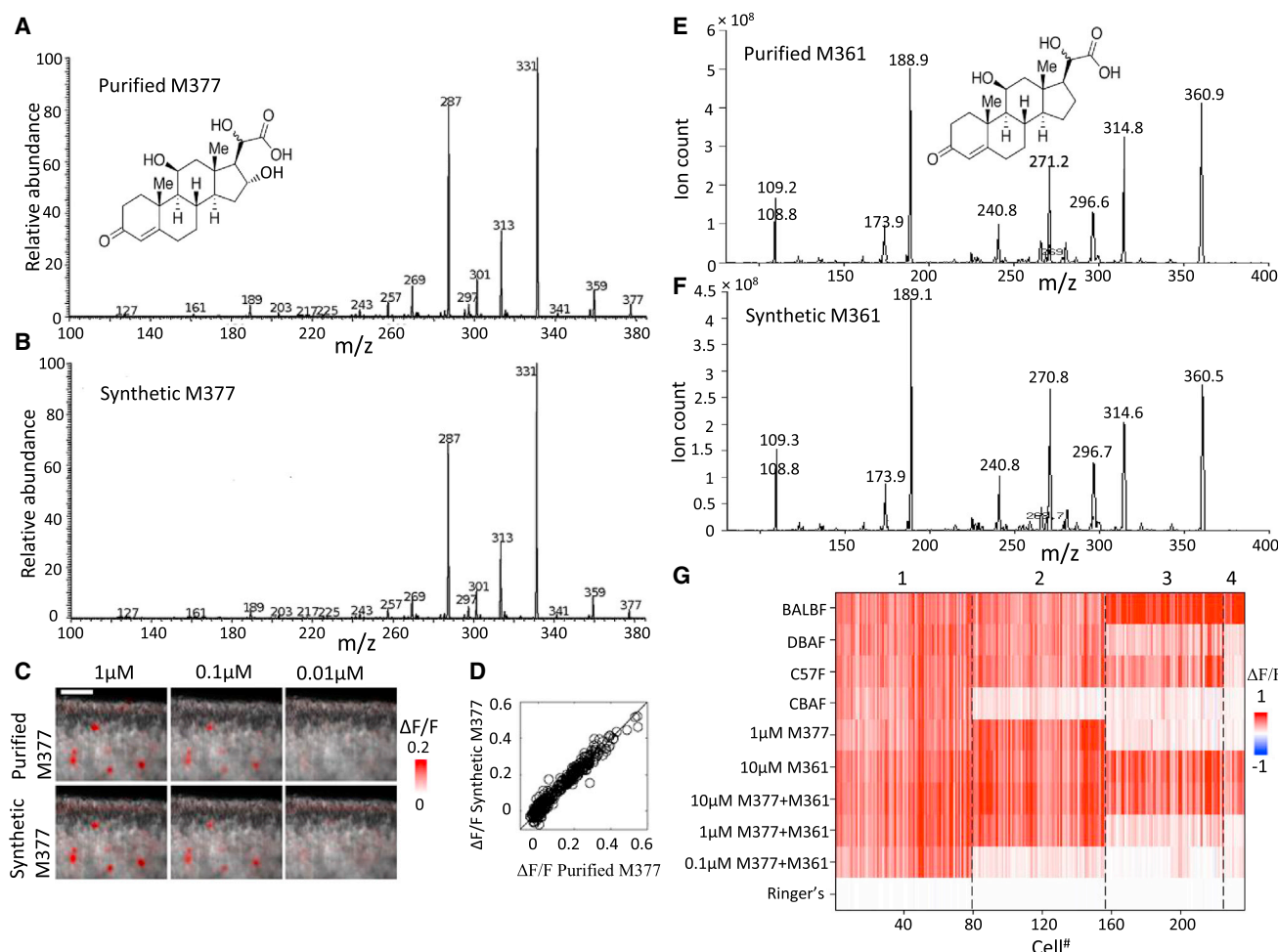


Figure 6. Structural Validation of M377 and M361

(A) Proposed chemical structure of purified M377 and its MS² product-ion spectrum.

(B) Structure of synthetic cortigynic acid and its MS² spectrum.

(C) Calcium imaging to purified and synthetic M377 over a range of matched concentrations.

(D) Response amplitude of each neuron to purified and synthetic M377, pooled across all concentrations. The average response across four trials by each neuron to each matched-concentration stimulus combination is shown as a circle.

(E and F) Chemical structure of purified and synthetic M361 and the MS² product-ion spectrum.

(G) Combinatorial coding of female-specific cues. Neuronal responses to 100-fold dilute urine stimuli, M377, M361, and their mixture at three different concentrations.

See also [Figures S5](#) and [S6](#) and [Table S2](#).

[Figure S3](#)), it seems likely that the partial strain-selectivity of this class of neurons stems from a threshold effect at the tested concentrations. The two other clusters responded to a subset of strains and were not activated by M377, but were activated by M361. Consequently, multiple classes of VSNs (likely expressing different vomeronasal receptor genes) participate in a combinatorial code for the presence and concentration of these two ligands ([Figure 6G](#)).

Cortigynic Acid Is Specific to Sexually Intact Adult Female Mice

Cortigynic acid is expressed in urine of all tested adult female strains at a concentration near 1 μM and is essentially absent

from the urine of adult males ([Figure S3](#)). To further explore the factors that control its expression, we collected urine from estrus and non-estrus females, ovariectomized females, and juvenile males and females (see [Experimental Procedures](#)). Quantitative LC-MS/MS of these samples showed that cortigynic acid was 7-fold less-concentrated in the urine of P21 juvenile females and more than 30-fold lower in the urine of ovariectomized females (data not shown). There was no cortigynic acid detected in urine from adrenalectomized males and females (data not shown), indicating that its synthesis is affected by both the gonads and the hypothalamic/pituitary/adrenal axis. Somewhat surprisingly, the concentration of cortigynic acid was only modestly (at most ~30%) modulated by the estrous cycle. To determine whether

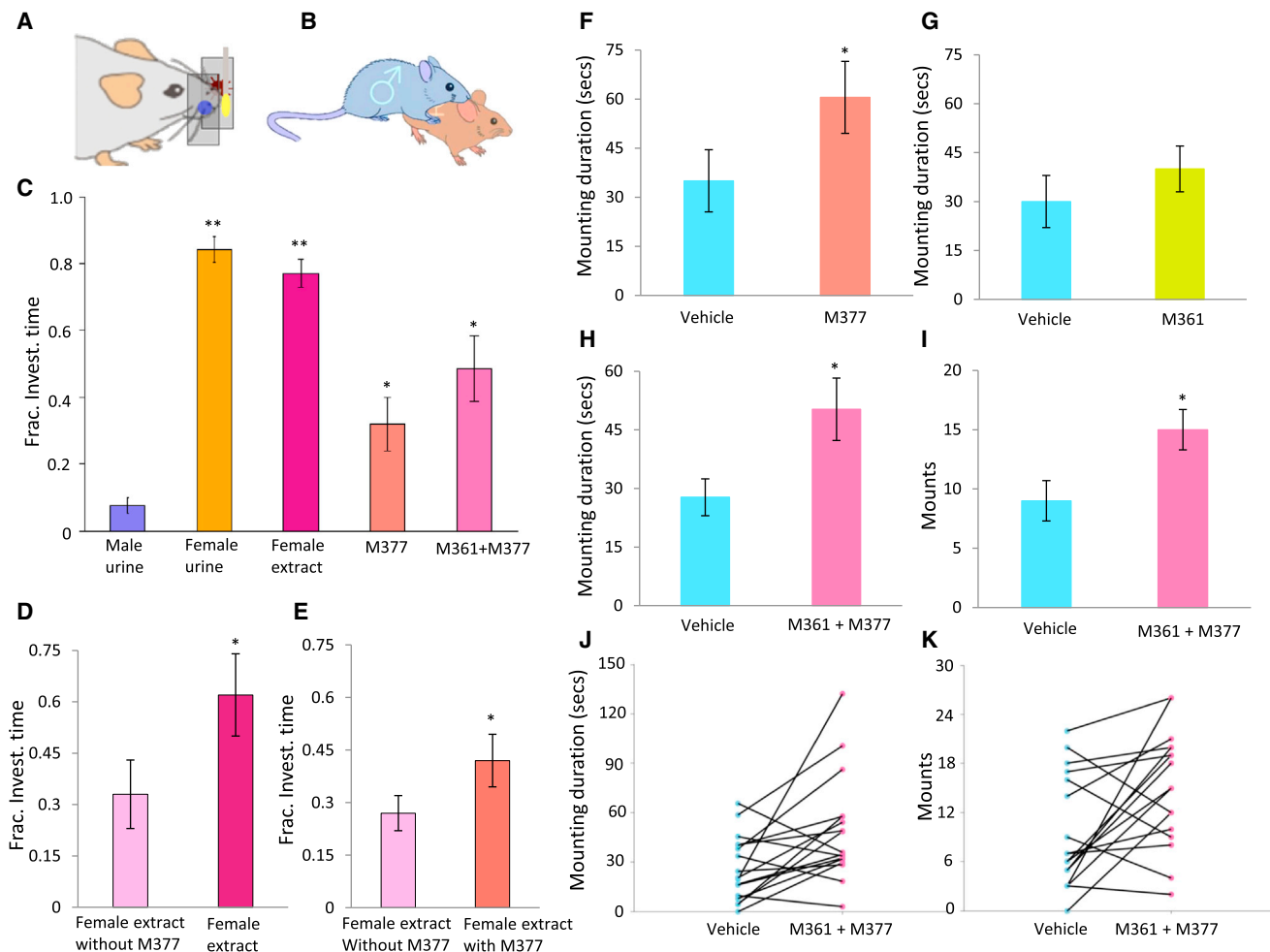


Figure 7. Male Investigatory and Mounting Behavior

(A and B) Schematic of the experiment. Male “sniff time” to stimuli presented on a cotton swab is measured by an optical beam-break detector; male mounting behavior was scored manually.

(C) Mean and SEM of the fraction of time spent investigating the swab after initial contact. Males spent longer investigating the sample doped with M377, M377 + M361, female extract, or female urine (* $p < 0.001$, ** $p < 0.00001$, $n = 11$, t test).

(D) Removing M377 from female urine extract significantly reduced wild-type males’ investigation time (* $p < 0.01$, $n = 12$, t test).

(E) Adding M377 back to the “deficient” female urine extract robustly increased male’s investigation time (* $p < 0.05$, $n = 12$, t test).

(F) Duration of mounting toward females painted with vehicle (light blue) and M377 (salmon) (* $p < 0.05$, $n = 12$, Mann-Whitney test).

(G) Duration of mounting toward females painted with vehicle (light blue) and M361 (lime).

(H) Duration of mounting toward females painted with vehicle (light blue) and M361 + M377 (pink) (* $p < 0.02$, $n = 17$, Mann-Whitney test).

(I) Frequency of mounting toward females painted with vehicle (light blue) and M361 + M377 (pink) (* $p < 0.02$, $n = 17$, Mann-Whitney test).

(J) Mounting duration of individual males toward females painted with vehicle (light blue) and M377 + M361 (pink).

(K) Mounting frequency of individual males toward females painted with vehicle (light blue) and M377 + M361 (pink).

See also Figure S7.

cortigynic acid is species specific, we also examined hamster and rabbit urine from both sexes. No cortigynic acid was detected in these two species, suggesting that cortigynic acid is a female-specific cue for mice.

Cortigynic Acid Increases Investigatory Behavior of Male Mice

Male mice investigate urine from females significantly longer than urine from males, in a manner that depends on direct phys-

ical contact and an intact vomeronasal system (Pankevich et al., 2004). This behavior depends on a positive cue in female urine, rather than a negative cue in male mouse urine: mixtures of the two stimuli cause similar investigatory behavior to that triggered by female urine alone, when presented at matching concentration (Guo and Holy, 2007).

We used an optical beam-break detector to record bouts of voluntary investigation by untrained, freely behaving male mice (Figure 7A). Because cortigynic acid, corticosteronic acid, and

female urine extract are not volatile and therefore attract little notice by themselves, we separately added them to whole male mouse urine. We compared investigatory behavior to a swab with just male mouse urine against one with male mouse urine doped with 10 μ M cortigynic acid, cortigynic acid + corticosteronic acid, female urine extract, and female urine. Each male was presented with each swab, plus a blank in between, in sessions of 120 s, spaced over a period of approximately 1 hr. The order of presentation of the stimuli was counterbalanced across mice, and the experimenter was blinded to the identity of the stimuli.

Males investigated female mouse urine for durations that were many times longer than exhibited for male mouse urine. Extracts of female-mouse urine were not significantly different ($p > 0.1$) from whole female mouse urine, demonstrating that their particular interest in this stimulus is largely triggered by the non-volatile components (Figure 7C; $**p < 0.00001$, $n = 11$).

The addition of cortigynic acid alone to male mouse urine also triggered a large increase in behavioral investigation (Figure 7C; $*p < 0.001$, $n = 11$). The addition of corticosteronic acid to cortigynic acid did not significantly increase the male investigation time compared with that caused by M377 on its own (Figure 7C; $p > 0.1$, $n = 11$). To determine whether this effect is simply a consequence of novelty, we also tested whether a structurally similar compound 1,4-pregnandien-11 β , 17, 20-triol-3-one-21-carboxylic acid (P0750) exhibited a similar effect. Using the same testing procedure, we found no difference between male urine and male urine doped with P0750 (Figure S7C; $p > 0.7$, $n = 10$), indicating that increased investigation is not a general reaction to any steroid carboxylic acid.

Transient receptor potential channels (TRPC2) play a fundamental role in detecting pheromonal and other chemical social recognition signals (Stowers et al., 2002; Leybold et al., 2002; Omura and Mombaerts, 2014). In *TrpC2*^{-/-} mice, cortigynic acid did not increase the investigation time (Figure S7D).

To complement the gain of function observed when adding cortigynic acid to male mouse urine, we asked whether cortigynic acid is necessary by removing it from female mouse urine. Female urine extract was separated by HPLC using a methanol/water gradient and then recombined, omitting the two fractions (out of 100) containing M377 (Figure S7B). In behavioral testing, male mice significantly decreased their investigation of the “deficient” female urine extract compared to the whole female urine extract (Figure 7D; $*p < 0.01$, $n = 12$). Adding cortigynic acid back to the “deficient” female urine extract significantly increased male investigation time (Figure 7E; $*p < 0.05$, $n = 12$). These results demonstrate that cortigynic acid plays a major and specific role in attracting male investigation.

Cortigynic Acid and Corticosteronic Acid Are Sufficient to Promote Mounting Behavior

Classic experiments demonstrated that female mouse urine contains one or more pheromones that stimulate male mounting behavior (Dixon and Mackintosh, 1975; Haga-Yamanaka et al., 2014); however, the chemical identity of the cue(s) has never been established. To determine whether cortigynic acid and corticosteronic acid might be the missing cue(s), we painted the anogenital area of ovariectomized females with these com-

pounds and exposed them to sexually naive males. Cortigynic acid alone significantly increased the total amount of time spent mounting (Figure 7F; $p < 0.05$, $n = 12$), but exhibited only a trend-level increase on the total number of mounting episodes (data not shown). In contrast, females painted with corticosteronic acid alone did not undergo significantly more mounting than those painted with water (Figure 7G; $p > 0.05$, $n = 12$). Interestingly, the combination of cortigynic acid and corticosteronic acid induced the most robust mounting activity (Figures 7H–7K; $p < 0.02$, $n = 17$). Thus, individually cortigynic acid is sufficient to induce mounting, but it exhibits a synergistic effect with corticosteronic acid.

DISCUSSION

To elucidate the vomeronasal code for identity, we developed CAM to perform a broad screen for ligands that “explain” the response of VSNs. This strategy exploits natural diversity to identify many candidate ligands from a single round of bioassays and is intrinsically parallelizable to multiple sensory neuron types. CAM’s only requirements—functionally diverse samples and a richly quantifiable bioassay—may be met in other studies, making the technique broadly applicable.

Without requiring any purification, CAM identified a single compound, out of 1,634 possibilities, to explain responses of a neuronal type selective for urine from BALB/c females (Figure 3). This single compound had been previously identified from BALB/c female-mouse urine (Nodari et al., 2008), and this finding demonstrated that CAM can identify a small handful of plausible ligands from hundreds or thousands of candidates. We organized neuronal responses into eight reproducibly identified classes, of which seven had one or more CAM candidates. Explaining these responses would seemingly require a minimum of seven distinct ligands; our statistical analysis selected a set of 23 plausible candidates. It is also worth noting that some of these 23 candidate ligands apply to more than one class of neuron (Table S1). Indeed, our detailed analysis of the class 5 (Figures 2C and S1) candidate ligands (M377 and M361) demonstrated that four distinct classes of neurons responded to these two ligands. These results suggest that vomeronasal coding may be organized around many receptor genes that are tuned with differing specificities and sensitivities to a much smaller number of behaviorally relevant ligands. This interpretation is consistent with the much larger number of receptor genes (~ 300) than candidate ligands (23) in urine. Moreover, the structural similarity of cortigynic acid and corticosteronic acid, like that found in an earlier ligand discovery (Nodari et al., 2008), suggests an emerging general pattern: that the vomeronasal system detects multiple intermediates in the same metabolic pathway, using multiple receptor genes as a combinatorial code (Figure 6G) to provide robust information about concentration and identity (Arnsen and Holy, 2013).

Neuronal responses to mouse urine were previously demonstrated to exhibit sex selectivity (He et al., 2008; Holy et al., 2000; Tolokh et al., 2013). However, no urinary ligand exhibiting a pattern of sex-, but not strain-, specific distribution had yet been reported. Cortigynic acid and corticosteronic acid are the first potential cues for encoding “sex” from urinary cues for the

mouse vomeronasal system. Cortigynic acid was present at high concentration only in intact adult female mice urine and was relatively constant across strains (Figure S3).

In behavioral tests, cortigynic acid substantially and significantly increased male investigatory and mounting behavior. When added to male urine, this one ligand increased the duration of investigation bouts by more than 2-fold. Conversely, removing cortigynic acid from female urine extract dramatically reduced male investigatory time (Figure 7D). Most remarkably, painting ovariectomized females with M377, alone or especially in conjunction with M361, significantly increased male mounting behavior. Based on results with a close structural analog, we speculate that the specific interest in this compound integrates multiple structural features, including the 16-hydroxyl functional group specific to female mice.

Previous results (Dixon and Mackintosh, 1975; Haga-Yamamoto et al., 2014) demonstrated that estrous urine, but not diestrous urine, was effective in promoting mounting behavior toward ovariectomized females. Here, we observed a robust increase in mounting with compounds whose concentration was at most modestly (~30%) influenced by the estrous cycle. There appear to be several possible interpretations of this result. First, it is possible that even a modest modulation of concentration is sufficient to lead to estrous-dependent behavior. Second, it is possible that estrous is signaled by the presence of cortigynic/corticosterone acid together with the absence of one or more (currently unknown) diestrous-specific cue(s). In such a scenario, the presence of cortigynic/corticosterone acid would mimic estrous even though these compounds do not exhibit dramatic cycling. Third, to ensure that we would not miss a positive effect, in mounting assays cortigynic acid and corticosterone acid were used at approximately 10-fold excess concentration compared to their natural concentrations; it is possible that these ligands cross-react to a receptor that is more sensitively tuned to a different, estrous-specific ligand. The conditions under which our urine samples were collected, with five animals per cage, were not optimal for studying the role of estrous; we anticipate that this will be an important topic for future studies.

Together with the identification of receptors, the identification of VSN ligands establishes a molecular foundation for understanding how sensory inputs are used in recognition and behavior. Although receptors tend to fall into gene families and can be exhaustively mapped via genome sequencing and homology analysis, organic ligands are diverse and have heretofore been discovered idiosyncratically. In this study, we developed a robust method that enabled us to systematically pick out plausible ligands for vomeronasal neurons. These results take a major step toward a comprehensive understanding of the molecular code for identity in the vomeronasal system.

EXPERIMENTAL PROCEDURES

Mice

Male B6D2F1 mice 8–12 weeks of age were used in all electrophysiology experiments. *tetO-GCaMP2/OMP-IRES-tTA* mice (He et al., 2008) and *Ai38(GCaMP3)/OMP-Cre* male mice aged 8–12 weeks were used for imaging by objective-coupled planar illumination (OCPI) microscopy (Holekamp et al., 2008). Prior to dissection, mice were anesthetized with carbon dioxide and decapitated. All procedures were approved by the Washington University

Animal Studies Committee. Sexually naive 3-month-old male B6D2F1 mice were used to examine mounting behaviors. Ovariectomy of 2-month-old C57BL/6J females was performed by a veterinary technician from the Department of Comparative Medicine at Washington University in St. Louis.

Electrophysiological Recording and Calcium Imaging

Dissection and recording procedures were performed as previously described (Amson et al., 2010; Turaga and Holy, 2012; Xu and Holy, 2013); briefly, intact B6D2F1 male vomeronasal epithelia were isolated, mounted on a multi-electrode array for electrophysiological recording, and adhered onto a nitrocellulose membrane (0.45 μ m; Millipore) for calcium imaging. Before recording, the VNO was acclimated by superfusing with Ringer's for at least 30 min. The order of stimuli, together with Ringer's control, was randomized across multiple repeated trials. Further details about stimulus preparation, recording method, and data analysis can be found in Supplemental Experimental Procedures.

Component-Activity Matching, Purification, Structural Identification, and Removal of M377 from Urine Extracts

See Supplemental Experimental Procedures.

Investigation Behavior

Male mouse chemosensory investigation was recorded as previously described (Guo and Holy, 2007; Holy and Guo, 2005). Periods during which the swab was investigated were extracted algorithmically from the voltage signal from an optical beam-break detector. Significance was assessed by a paired *t* test. Further detail is presented in Supplemental Experimental Procedures.

Mounting Behavior

Individual B6D2F1 males were singly housed for at least 2 weeks without bedding change prior to the mounting test. Assays were performed in their home cages. Videos were scored manually using Interact Mangold software (Mangold International). Statistical significance was assessed by the Mann-Whitney test. Further detail is presented in Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.012>.

AUTHOR CONTRIBUTIONS

X.F. and T.E.H. designed the project. X.F. performed the electrophysiology experiments and analysis. I.G.-V. performed LC-MS/MS profiling. T.E.H. developed the CAM method and analyzed the data. X.F. purified M377 and M361. Y.Y., X.F., and M.L.G. performed the structural analysis. P.S.X. and X.F. performed calcium imaging and analysis of data. X.F. and W.C. performed the behavior tests. T.E.H. and X.F. did the analysis. X.F. and T.E.H. wrote the manuscript with contributions from all authors.

ACKNOWLEDGMENTS

We thank Amber Tyler for helping collect estrus female mice urine and Petland St. Louis (south city store) for helping collect hamster and rabbit urine. We thank members of the T.E.H. and M.L.G. labs, as well as the anonymous referees, for many constructive comments. This work was supported by NIH/NIDCD (R01 DC005964), NIH/NIMS/NIAAA (R01 NS068409), and NIH/NIDCD (R01 DC010381) (to T.E.H.) and NIH/NIGMS (P41 GM103422-36) (to M.L.G.).

Received: May 19, 2015

Revised: July 30, 2015

Accepted: August 6, 2015

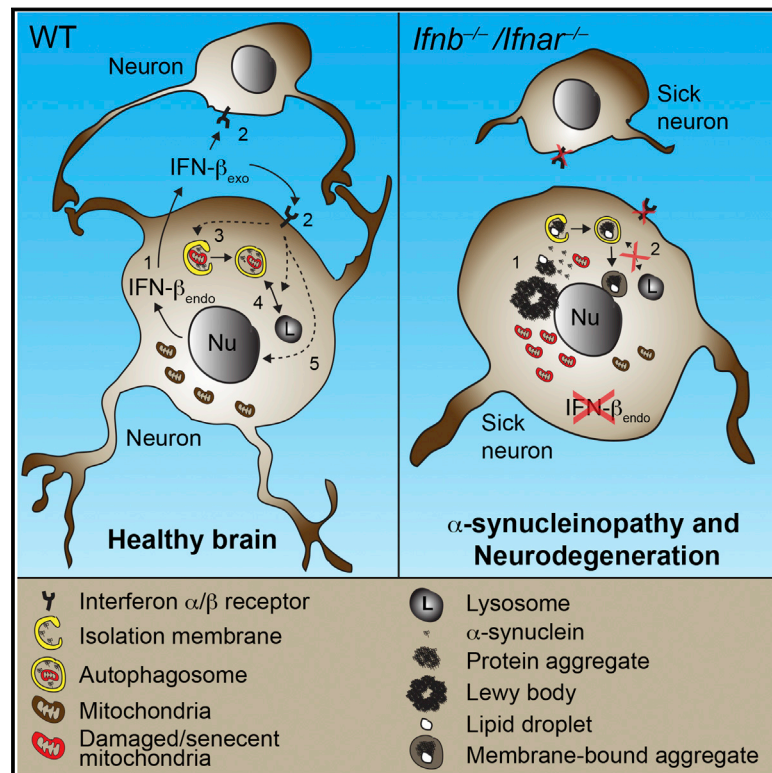
Published: October 1, 2015

REFERENCES

- Amson, H.A., and Holy, T.E. (2013). Robust encoding of stimulus identity and concentration in the accessory olfactory system. *J. Neurosci.* 33, 13388–13397.
- Amson, H.A., Fu, X., and Holy, T.E. (2010). Multielectrode array recordings of the vomeronasal epithelium. *J. Vis. Exp.*, 2010 Mar 1;(37). pii: 1845.
- Brennan, P.A., and Keverne, E.B. (1997). Neural mechanisms of mammalian olfactory learning. *Prog. Neurobiol.* 51, 457–481.
- Chamero, P., Marton, T.F., Logan, D.W., Flanagan, K., Cruz, J.R., Saghatelian, A., Cravatt, B.F., and Stowers, L. (2007). Identification of protein pheromones that promote aggressive behaviour. *Nature* 450, 899–902.
- Dixon, A.K., and Mackintosh, J.H. (1975). The relationship between the physiological condition of female mice and the effects of their urine on the social behaviour of adult males. *Anim. Behav.* 23, 513–520.
- Dulac, C., and Torello, A.T. (2003). Molecular detection of pheromone signals in mammals: from genes to behaviour. *Nat. Rev. Neurosci.* 4, 551–562.
- Guo, Z., and Holy, T.E. (2007). Sex selectivity of mouse ultrasonic songs. *Chem. Senses* 32, 463–473.
- Haga-Yamanaka, S., Ma, L., He, J., Qiu, Q., Lavis, L.D., Looger, L.L., and Yu, C.R. (2014). Integrated action of pheromone signals in promoting courtship behavior in male mice. *eLife* 3, e03025.
- Halpern, M., and Martínez-Marcos, A. (2003). Structure and function of the vomeronasal system: an update. *Prog. Neurobiol.* 70, 245–318.
- He, J., Ma, L., Kim, S., Nakai, J., and Yu, C.R. (2008). Encoding gender and individual information in the mouse vomeronasal organ. *Science* 320, 535–538.
- Hill, A.V. (1938). The heat of shortening and dynamics constants of muscles. *Proc. R. Soc. Lond. B* 126, 136–195.
- Holekamp, T.F., Turaga, D., and Holy, T.E. (2008). Fast three-dimensional fluorescence imaging of activity in neural populations by objective-coupled planar illumination microscopy. *Neuron* 57, 661–672.
- Holy, T.E., and Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biol.* 3, e386.
- Holy, T.E., Dulac, C., and Meister, M. (2000). Responses of vomeronasal neurons to natural stimuli. *Science* 289, 1569–1572.
- Hsu, F.F., Nodari, F., Kao, L.F., Fu, X., Holekamp, T.F., Turk, J., and Holy, T.E. (2008). Structural characterization of sulfated steroids that activate mouse pheromone-sensing neurons. *Biochemistry* 47, 14009–14019.
- Kimchi, T., Xu, J., and Dulac, C. (2007). A functional circuit underlying male sexual behaviour in the female mouse brain. *Nature* 448, 1009–1014.
- Kimoto, H., Haga, S., Sato, K., and Touhara, K. (2005). Sex-specific peptides from exocrine glands stimulate mouse vomeronasal sensory neurons. *Nature* 437, 898–901.
- Leypold, B.G., Yu, C.R., Leinders-Zufall, T., Kim, M.M., Zufall, F., and Axel, R. (2002). Altered sexual and social behaviors in *trp2* mutant mice. *Proc. Natl. Acad. Sci. USA* 99, 6376–6381.
- Ma, M.H. (2007). Encoding olfactory signals via multiple chemosensory systems. *biochemistry and molecular biology. Biochemistry and molecular biology* 42, 463–480.
- Nodari, F., Hsu, F.F., Fu, X., Holekamp, T.F., Kao, L.F., Turk, J., and Holy, T.E. (2008). Sulfated steroids as natural ligands of mouse pheromone-sensing neurons. *J. Neurosci.* 28, 6407–6418.
- Omura, M., and Mombaerts, P. (2014). *Trpc2*-expressing sensory neurons in the main olfactory epithelium of the mouse. *Cell Rep.* 8, 583–595.
- Pankevich, D.E., Baum, M.J., and Cherry, J.A. (2004). Olfactory sex discrimination persists, whereas the preference for urinary odorants from estrous females disappears in male mice after vomeronasal organ removal. *J. Neurosci.* 24, 9451–9457.
- Raju, U., Bradlow, H.L., and Levitz, M. (1990). Estriol-3-sulfate in human breast cyst fluid. Concentrations, possible origin, and physiologic implications. *Ann. N Y Acad. Sci.* 586, 83–87.
- Stowers, L., Holy, T.E., Meister, M., Dulac, C., and Koentges, G. (2002). Loss of sex discrimination and male-male aggression in mice deficient for *TRP2*. *Science* 295, 1493–1500.
- Tolokh, I.I., Fu, X., and Holy, T.E. (2013). Reliable sex and strain discrimination in the mouse vomeronasal organ and accessory olfactory bulb. *J. Neurosci.* 33, 13903–13913.
- Turaga, D., and Holy, T.E. (2012). Organization of vomeronasal sensory coding revealed by fast volumetric calcium imaging. *J. Neurosci.* 32, 1612–1621.
- Wang, T., Shah, Y.M., Matsubara, T., Zhen, Y., Tanabe, T., Nagano, T., Fotso, S., Krausz, K.W., Zabriskie, T.M., Idle, J.R., and Gonzalez, F.J. (2010). Control of steroid 21-oic acid synthesis by peroxisome proliferator-activated receptor alpha and role of the hypothalamic-pituitary-adrenal axis. *J. Biol. Chem.* 285, 7670–7685.
- Xu, P.S., and Holy, T.E. (2013). Whole-mount imaging of responses in mouse vomeronasal neurons. *Methods Mol. Biol.* 1068, 201–210.

Lack of Neuronal IFN- β -IFNAR Causes Lewy Body- and Parkinson's Disease-like Dementia

Graphical Abstract



Authors

Patrick Ejlerskov,
Jeanette Göransdotter Hultberg,
JunYang Wang, ..., Marco Prinz,
David C. Rubinsztein,
Shohreh Issazadeh-Navikas

Correspondence

shohreh.issazadeh@bric.ku.dk

In Brief

Lack of the immunomodulatory cytokine interferon- β (IFN- β) causes spontaneous neurodegeneration resembling sporadic Lewy body and Parkinson's disease dementia due to defects in neuronal autophagy.

Highlights

- Lack of neuronal IFN- β -IFNAR signaling causes brain Lewy body accumulation
- IFN- β deficiency causes late-stage autophagy block and thereby α -synuclein aggregation
- IFN- β promotes neuronal autophagy and α -synuclein clearance
- *Ifnb* gene therapy prevents dopaminergic neuron loss in a familial PD model

Accession Numbers

GSE63815



Lack of Neuronal IFN- β -IFNAR Causes Lewy Body- and Parkinson's Disease-like Dementia

Patrick Ejlerskov,¹ Jeanette Göransdotter Hultberg,¹ JunYang Wang,¹ Robert Carlsson,¹ Malene Ambjørn,¹ Martin Kuss,¹ Yawei Liu,¹ Giovanna Porcu,¹ Kateryna Kolkova,¹ Carsten Friis Rundsten,¹ Karsten Ruscher,² Bente Pakkenberg,³ Tobias Goldmann,⁴ Desiree Loreth,⁵ Marco Prinz,^{4,6} David C. Rubinsztein,⁷ and Shohreh Issazadeh-Navikas^{1,*}

¹Biotech Research and Innovation Centre, University of Copenhagen, 2200 Copenhagen, Denmark

²Department of Clinical Sciences, Lund University, 22100 Lund, Sweden

³Research Laboratory for Stereology and Neuroscience, Bispebjerg University Hospital, 2200 Copenhagen, Denmark

⁴Institute for Neuropathology

⁵Department of Neurology

⁶Centre for Biological Signaling Studies

University of Freiburg, 79106 Freiburg, Germany

⁷Department of Medical Genetics, Cambridge Institute for Medical Research, Cambridge CB2 0XY, UK

*Correspondence: shohreh.issazadeh@bric.ku.dk

<http://dx.doi.org/10.1016/j.cell.2015.08.069>

SUMMARY

Neurodegenerative diseases have been linked to inflammation, but whether altered immunomodulation plays a causative role in neurodegeneration is not clear. We show that lack of cytokine interferon- β (IFN- β) signaling causes spontaneous neurodegeneration in the absence of neurodegenerative disease-causing mutant proteins. Mice lacking *Ifnb* function exhibited motor and cognitive learning impairments with accompanying α -synuclein-containing Lewy bodies in the brain, as well as a reduction in dopaminergic neurons and defective dopamine signaling in the nigrostriatal region. Lack of IFN- β signaling caused defects in neuronal autophagy prior to α -synucleinopathy, which was associated with accumulation of senescent mitochondria. Recombinant IFN- β promoted neurite growth and branching, autophagy flux, and α -synuclein degradation in neurons. In addition, lentiviral IFN- β overexpression prevented dopaminergic neuron loss in a familial Parkinson's disease model. These results indicate a protective role for IFN- β in neuronal homeostasis and validate *Ifnb* mutant mice as a model for sporadic Lewy body and Parkinson's disease dementia.

INTRODUCTION

Neurodegenerative diseases have disrupted neuronal homeostasis and their pathologies often overlap. Protein aggregates containing α -synuclein (α -syn), which eventually forms larger Lewy bodies (LBs), are seen in Parkinson's disease (PD), dementia with Lewy bodies (DLB), multiple system atrophy, and in some forms of Alzheimer's disease (AD), all neurodegenerative diseases associated with aging (Arima et al., 1999; Francis, 2009;

Lippa et al., 1998). Often, the protein aggregates contain hyperphosphorylated tau and ubiquitin (Jellinger and Attems, 2008).

Neurodegenerative events in these diseases are linked to inflammation (Mrak and Griffin, 2007; Tansey et al., 2008), but, despite this link, defects in genes regulating inflammation do not have an established causative role in neurodegeneration. We report that deletion of *Ifnb*, which encodes interferon- β (IFN- β), is sufficient to cause a cascade of neurodegenerative events. IFN- β belongs to the type I interferon family of cytokines and binding to its receptor, interferon- α/β receptor (IFNAR), results in immunoregulation including anti-viral and anti-inflammatory effects (Prinz et al., 2008; Teige et al., 2003, 2006) and benefits for multiple sclerosis patients (Liu et al., 2014; Yong et al., 1998). However, the role of IFN- β in classic neurodegenerative diseases is unknown. We report that *Ifnb*^{-/-} mice developed spontaneous pathologies mimicking major aspects of human neurodegeneration such as PD and DLB. *Ifnb*^{-/-} mice had age-associated motor learning defects, neuromuscular deficiencies, and cognitive impairment. *Ifnb*^{-/-} pathology was associated with LBs resulting from defective neuronal autophagy. Autophagy, a pathway that degrades long-lived proteins, organelles, lipids, and protein aggregates, is essential for neuronal homeostasis (Harris and Rubinsztein, 2012), and deleting neural autophagy-regulating genes leads to neurodegeneration (Hara et al., 2006; Komatsu et al., 2006).

Our findings indicate a central role for IFN- β in neuronal homeostasis as a regulator of autophagy-mediated protein degradation and accentuates *Ifnb*^{-/-} mice as a model for neurodegenerative diseases with α -synucleinopathy and dementia such as PD and DLB.

RESULTS

Ifnb^{-/-} Mice Exhibit Behavioral and Cognitive Impairments and Neurodegeneration

We examined the effect of *Ifnb* gene deletion on motor coordination, learning, and grip strength. *Ifnb*^{-/-} mice were significantly impaired in motor coordination and learning from 3 months

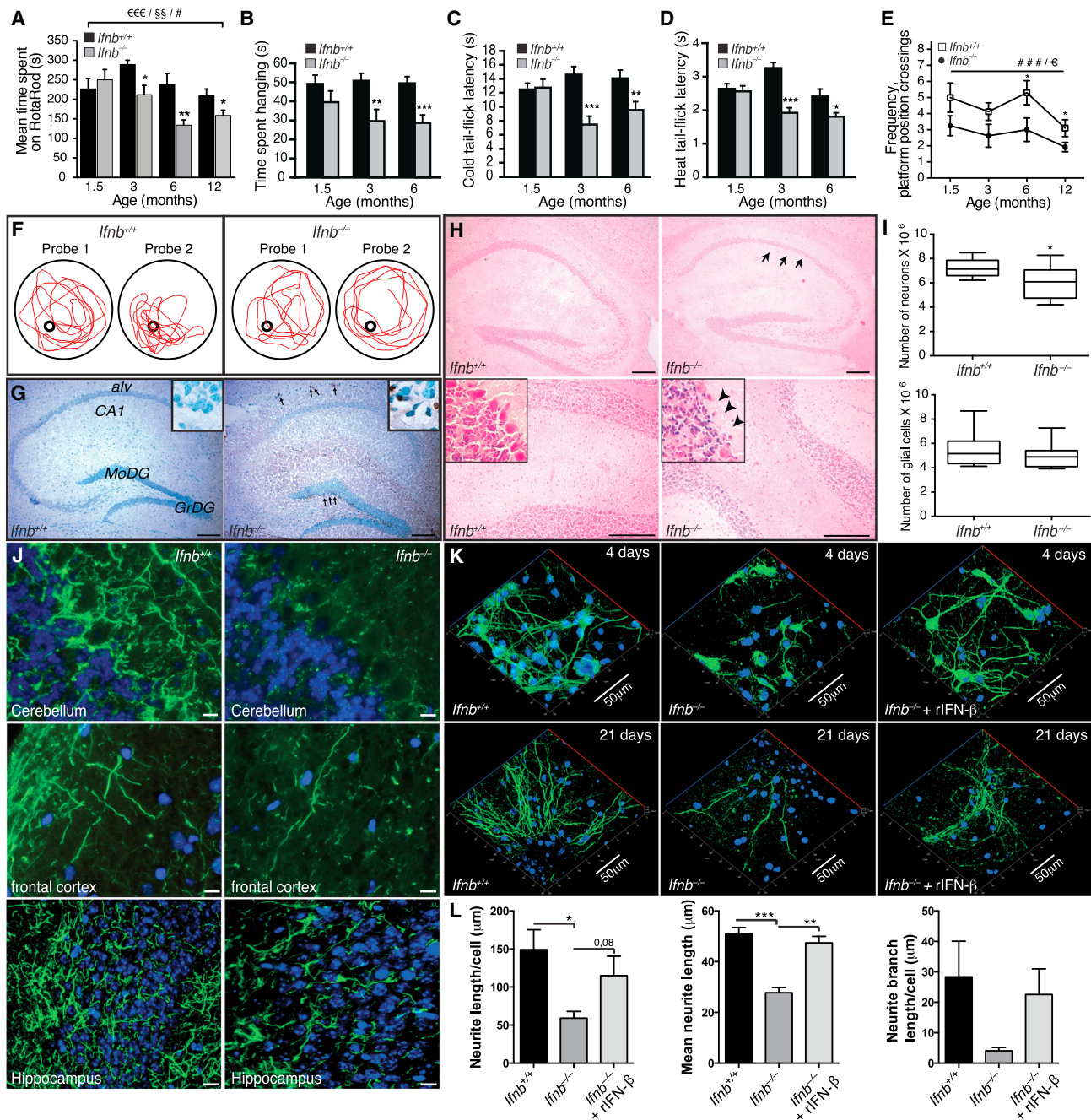


Figure 1. *Ifnb*^{-/-} Mice Exhibit Spontaneous Behavioral and Cognitive Impairments and Neurodegeneration

(A) Motor coordination by RotaRod. Data are mean second(s) \pm SEM, $n = 8-13$ /group. * $p < 0.05$, ** $p < 0.01$ for genotype effect per time point by unpaired Student's t test; # $p < 0.05$ for aging effect. €€€ $p < 0.001$ and §§ $p < 0.01$ for genotype and aging effect, respectively, over time, by two-way ANOVA.

(B) Wire suspension performance. Data are mean \pm SEM, 9–13/group.

(C and D) (C) Cold and (D) heat tail-flick latency. Data are mean \pm SEM, 9–11/group.

(B–D) Unpaired Student's t test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(E) Morris water maze test. Data are mean \pm SEM, 8–21 mice/group. ### $p < 0.001$ shows genotype effect; € $p < 0.05$ shows age effect; * $p < 0.05$ comparing genotypes by two-way ANOVA and (*) with Bonferroni post hoc test.

(F) Representative maze swimming pattern on probes 1 and 2.

(G) TUNEL staining of hippocampus from 1.5-month-old mice with methyl green nuclear counterstaining. MoDG, molecular dentate gyrus; GrDG, granular dentate gyrus; alv, alveus. Scale bar, 200 μ m.

(H) H&E brain staining showing loss of granular cell layer and Purkinje cells (arrows) in *Ifnb*^{-/-} hippocampus (upper) and cerebellum (lower). Scale bar, 200 μ m.

(legend continued on next page)

compared to age-, sex-, and weight-matched wild-type (WT) *lfnb*^{+/+} littermates and in latency-to-fall time in a wire-suspension test (Figures 1A and 1B). We saw no differences in 1.5-month-old mice, suggesting that impaired motor coordination, balance, and grip strength were associated with age. During repeated motor-learning trials, retention time improved significantly in older *lfnb*^{+/+} mice compared to *lfnb*^{-/-} mice from age 3 months (Figure S1A), showing that reduced motor-learning in *lfnb*^{-/-} mice was associated with aging.

We assessed somatosensory function with nociception cold- and heat-induced tail-flick tests. Latency to tail flick was significantly shorter in *lfnb*^{-/-} than *lfnb*^{+/+} mice (Figures 1C and 1D), indicating hyperalgesia and defective nociception toward temperature-induced pain.

Forced swimming tests found no differences between *lfnb*^{-/-} and *lfnb*^{+/+} mice in swimming pattern, climbing effort, or immobility (Figure S1), so *lfnb*^{-/-} mice were not defective in locomotor activity in water in contrast to land. In water maze tests, *lfnb*^{-/-} mice had significant spatial- and memory-learning deficits that increased with age. During second tests, 6- and 12-month-old, *lfnb*^{-/-} mice had significantly fewer platform position crossings compared to *lfnb*^{+/+} mice, indicating impaired reference memory (Figures 1E and 1F). *lfnb*^{+/+} mice improved between the first and second tests at all ages except 12 months; *lfnb*^{-/-} mice did not improve, indicating impaired spatial learning. This was seen at all ages in *lfnb*^{-/-} mice measured by escape latency time in the second learning block (Figures S1B and S1C). Thus, although *lfnb*^{-/-} mice began with no major behavioral and cognitive defects, they developed age-dependent deficits with increased penetrance (Table S1).

IFN- β Is Essential for Neuronal Survival, Neurite Outgrowth, and Branching

We investigated whether *lfnb*^{-/-} behavioral deficits were associated with neurodegeneration. Apoptotic cells were detected in 1.5-month-old *lfnb*^{-/-} granular layers of olfactory bulbs (Figure S1E), the granular dentate gyrus of hippocampus and the subventricular zone (Figure 1G), and the striatum (STR) caudate putamen including the ependymal cell layer in 12-month-old *lfnb*^{-/-} mice (Figure S1F), but not detected in *lfnb*^{+/+} sagittal brain sections at similar ages. Apoptotic neurons increased in cultured *lfnb*^{-/-} primary cerebellar granular neurons (CGNs) and was reversed by recombinant (r)IFN- β (Figures S1G and S1H). Neurons were significantly reduced in the hippocampal CA1 region in 3- to 6-month-old *lfnb*^{-/-} mice and decreased in Purkinje cells of cerebellum. Glial cell counts were unchanged (Figures 1H and 1I).

Supporting in vivo deficits in neuronal circuits, *lfnb*^{-/-} mice had reduced neurite network formation in the cerebellum, frontal cortex, and hippocampus granular cell layer (Figure 1J). This was confirmed in cultured *lfnb*^{-/-} primary cortical neurons (CNs) that had reduced neurite length per cell, mean length, and

branch length per cell after 4 days, with restoration by rIFN- β (Figures 1K and 1L). A similar trend was seen after 21 days of culture (Figure 1K). Cultured primary *lfnb*^{-/-} CGNs mimicked the findings of CNs (Figures S1I and S1J). Thus, behavioral defects in old *lfnb*^{-/-} mice were associated with neuron death and reduced neurite circuits.

lfnb^{-/-} Neuron Gene Profiling Reveals Neurodegenerative Paths

We identified neuron-specific signaling and disease pathways caused by IFN- β deficiency with expression microarrays on highly pure (>98%) primary CGNs from *lfnb*^{-/-} and *lfnb*^{+/+} mice plus or minus rIFN- β . *lfnb*^{-/-} neurons had 323 upregulated and 233 downregulated genes (Figures 2A, 2B, and S2A). rIFN- β treatment had differential effects on *lfnb*^{+/+} versus *lfnb*^{-/-} CGNs with approximately equal numbers of downregulated and upregulated genes, but with relatively few overlaps (Figures 2C–2F). Multiple genes in the *lfnb*^{-/-} data set annotated as neuronal degeneration and the top ten pathways included cell death and neurological disorders. Genes in neurite formation and branching suggested compromised neurogenesis pathways in *lfnb*^{-/-} neurons, supporting the in vitro and in vivo phenotypes (Figures 2G and 2H). To investigate whether *lfnb* and signaling via IFNAR affected neurogenesis in vivo, 3-month-old *lfnb*^{-/-} and IFNAR knockout (KO) mice (*lfnar*^{-/-}) and WT littermates were injected with bromodeoxyuridine (BrdU). Lack of *lfnb* or receptor was associated with reduced neurogenesis in the hippocampus; dentate gyrus (Figure 2I), supporting the microarray data. IFN- β might directly affect prosurvival mechanisms as neurotrophin was among the top hits in gene set enrichment analysis (GSEA) (Table S2) and mRNA for the corticogenesis-regulating transcription factor Tox (Artegiani et al., 2015) was reduced in *lfnb*^{-/-} neurons (Figure S2A). Increased apoptosis and reduced neurogenesis might contribute to impaired cognition and reduced hippocampal neuron numbers. GSEA of *lfnb*^{+/+} and *lfnb*^{-/-} showed enrichment of genes associated with Huntington's disease (HD), PD, AD, and prion diseases, and rIFN- β caused the *lfnb*^{-/-} profile to resemble WT (Figure 2J), supporting that common pathways link neurodegenerative diseases (Shulman and De Jager, 2009).

We compared the *lfnb*^{-/-} neuronal gene profile with PD and HD mouse models (Fossale et al., 2011). Three sets of differentially expressed genes were generated: differentially expressed in our data ($p < 0.001$), in our data and the HD model ($p < 0.01$), or our data and the PD model ($p < 0.01$) (Figures S2A–S2C; Data S1). Clustered genes in *lfnb*^{-/-} neurons and the HD model were both similarly and oppositely regulated. In contrast, most clustered genes in *lfnb*^{-/-} neurons and the PD model were regulated similarly, thus indicating a higher gene signature resemblance with PD than HD (Figures S2B and S2C).

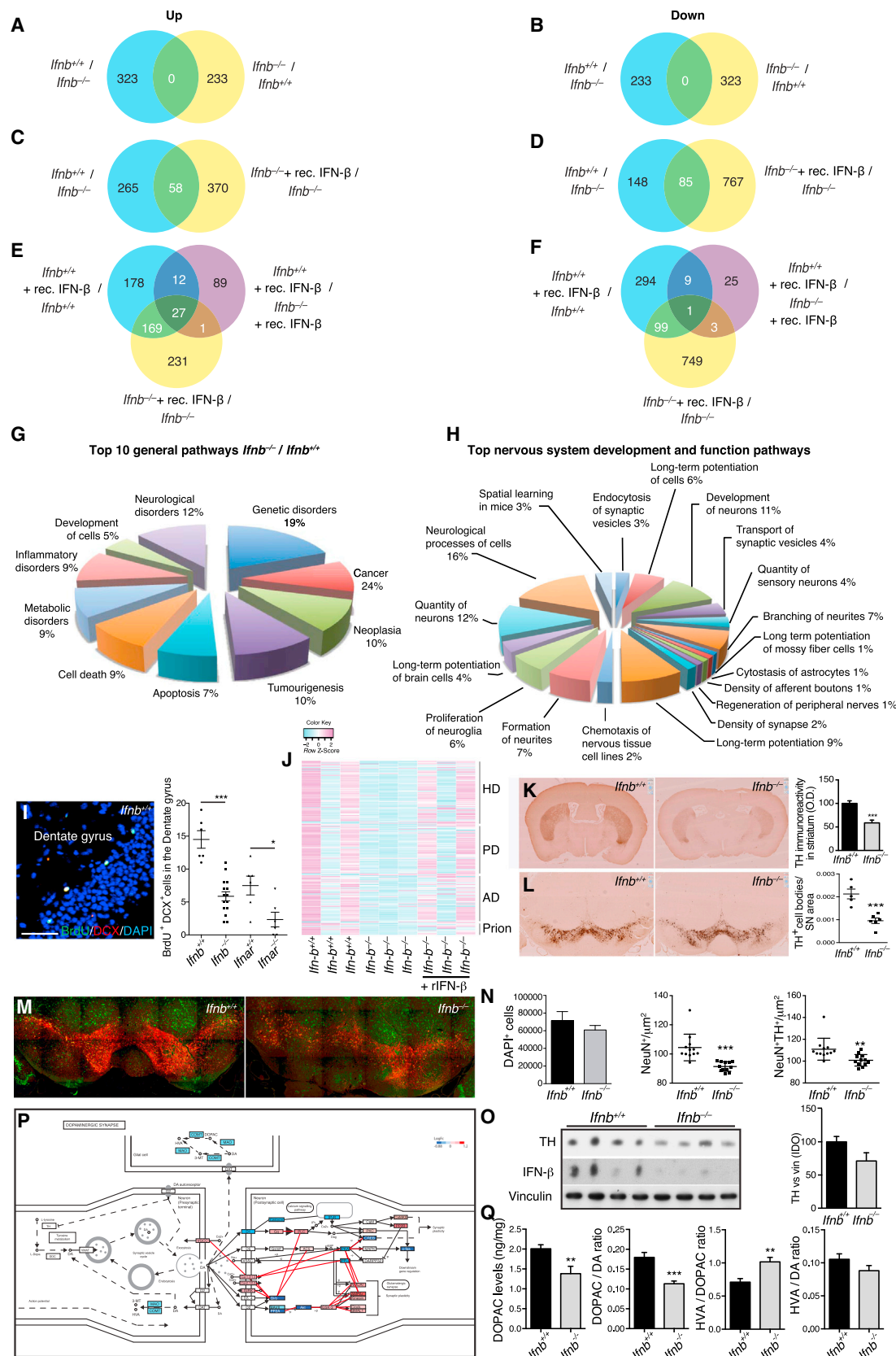
(I) Number of neurons and glial cells in hippocampus of 3- to 6-month-old mice. Data are mean \pm SEM, $n = 6$ mice/group. * $p < 0.05$ using Student's t test.

(J) IF of 3-month-old brain sections. Scale bars, 10 μ m.

(K) Z stack projections of CNs cultured for 4 and 21 days with or without rIFN- β . Scale bar, 50 μ m.

(L) Quantified neurite length from 4-day-old CNs.

Data are mean \pm SEM, $n = 3$ –4, counting 11–32 cells per experiment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by one-way ANOVA. See also Figure S1 and Table S1.



(legend on next page)

Lack of *Ifnb* Causes Defects in the Nigrostriatal Dopaminergic Pathway

Based on the possible PD-related defect in *Ifnb*^{-/-} neurons, we investigated nigrostriatal region integrity. Tyrosine hydroxylase (TH)⁺ fiber density and TH⁺ (dopamine-producing) neurons were significantly reduced in the STR and substantia nigra (SN) in *Ifnb*^{-/-} mice versus WT (Figures 2K and 2L). NeuN⁺ and NeuN⁺TH⁺ cells were reduced in the ventral midbrain, which was correlated with reduced TH protein in basal ganglia (BG) in *Ifnb*^{-/-} mice while total cells were unaffected (Figures 2M–2O). Among differentially expressed genes in *Ifnb*^{-/-} neurons, 31 ($p = 6.937 \times 10^{-4}$) were involved in regulating dopamine (DA) signaling (Figure 2P), which is involved in coordinating movements. DA signaling was also defective in the STR. No major differences were found in levels of DA or its metabolite homovanillic acid (HVA) (data not shown) in 3- to 6-month-old *Ifnb*^{-/-} mice, but the DA metabolite dihydroxyphenylacetic acid (DOPAC) was significantly lower, thus significantly affecting DA/DOPAC and DOPAC/HVA ratios (Figure 2Q). These findings support the importance of IFN- β in regulating dopamine turnover and protecting dopaminergic neurons.

Lack of Neuronal IFN- β -IFNAR Signaling Causes Lewy Bodies

Since *Ifnb*^{-/-} mice defects overlapped with neurodegenerative diseases, particularly PD, we examined *Ifnb*^{-/-} brain pathology. Gross anatomy was unchanged, but histological examination showed that *Ifnb*^{-/-} neuron degeneration was associated with age-dependent α -synucleinopathy. Staining for α -syn was normal in 1.5-month-old *Ifnb*^{-/-} brains; by 3 months, α -syn was found in LB-like structures in SN; however, α -syn staining intensity was reduced in 12-month-old *Ifnb*^{-/-} mice (Figures S3A and S3B), likely reflecting degeneration of TH⁺ neuron (Figure 2L). Alpha-syn and large pathogenic aggregates of phosphorylated (pSer129) α -syn were found in TH⁺ DA neurons of SN (Figure 3A). At 3 months, α -syn aggregates were widespread in the STR, frontal cortex (Figure 3B), hippocampus, and cerebellum (Figures S3C and S3D). Alpha-syn⁺ aggregates and neurites were found sporadically in thalamus, the brainstem, and subthalamic regions of 3-month-old *Ifnb*^{-/-} mice (Figures S3E–S3G). Neurons with α -syn⁺ LB-like structures increased with age (6- and

12-month-old) in *Ifnb*^{-/-} mouse thalamus (Figures 3C and 3D). Whole-brain protein extracts from 1.5-month-old mice were not different from WT, but 3-month-old mice had significantly increased α -syn in insoluble fractions and no difference in soluble fractions (Figures 3E–3I). No difference in mRNA for α -syn was observed between *Ifnb*^{-/-} and WT brains (Figure 3J); thus, the α -syn accumulation was not due to increased transcription. The insoluble α -syn fraction in the BG (including SN) was significantly higher in older mice, but the soluble fraction was still unchanged. A significant increase in high-molecular-weight dimeric, trimeric, and oligomeric α -syn were seen in *Ifnb*^{-/-} mice, but with a decrease in tetramers. While aggregated α -syn oligomers are neurotoxic (Rockenstein et al., 2014), α -syn tetramers are suggested to be the normal aggregation resistant conformation of the protein (Bartels et al., 2011; Wang et al., 2011), possibly explaining the tetramer abundance in *Ifnb*^{+/-} mice. pSer129- α -syn, which is prone to form pathogenic fibrillar aggregates, was increased in *Ifnb*^{-/-} mice (Figures 3K–3M). *Ifnar*^{-/-} brains showed a similar pattern to *Ifnb*^{-/-} regarding α -syn and pSer129- α -syn accumulation (Figure S3H). Blocking with α -syn_{96–140} peptide confirmed specific immunoreactivity for α -syn in *Ifnb*^{-/-} brains. To ensure that lack of *Ifnb* did not generate crossreacting α -syn-independent aggregates, we generated *Scna*^{-/-}*Ifnb*^{-/-} double-KO mice (DKO). *Ifnb*^{-/-} but not *Scna*^{-/-} or DKO mice had α -syn⁺ aggregates or LBs (Figures S3I–S3K).

Accumulated polyubiquitinated proteins are associated with neurodegenerative disease (Davies et al., 1997; Hara et al., 2006; Komatsu et al., 2006). While 1.5-month-old WT and *Ifnb*^{-/-} mice showed no differences, 3-month-old *Ifnb*^{-/-} brains had increased polyubiquitin (Figures 3N–3R), in hippocampus and cerebellum (Figures S3L and S3M), SN and locus coeruleus (Figures 3O and 3P) and vestibular nuclei of pons (not shown). Ubiquitin accumulated with inclusion body abundance, as seen in the locus coeruleus of 12-month-old *Ifnb*^{-/-} mice (Figure 3Q) and α -syn⁺ aggregates colocalizing with ubiquitin and phosphorylated Tau (pTau) significantly increased (Figures 3S and 3T).

To exclude systemic immune response influences, we analyzed *Ifnar*^{-/-} and *nes*^{Cre}:*Ifnar*^{fl/fl} mice. Consistent with *Ifnb*^{-/-} mice, α -syn⁺ LB-like structures were found in the

Figure 2. Gene Profiling of *Ifnb*^{-/-} Neurons Shows Association with Neurodegenerative Diseases

Mouse Affymetrix 430 2.0 arrays.

(A–F) Venn diagrams for gene expression analysis of *Ifnb*^{-/-}, *Ifnb*^{+/-} CGNs with or without rIFN- β , $n = 3$. Genes differentially regulated by 1.5-fold, $*p < 0.05$. (A, C, and E) Upregulated. (B, D, and F) downregulated genes.

(G and H) Gene expression analyzed for (G) top ten general signaling pathways and (H) nervous system development and function pathways.

(I) Neurogenesis in hippocampus dentate gyrus showing IF of an *Ifnb*^{+/-} mouse and quantifying all groups: $n = 4$ mice/group and six to 12 sagittal sections were counted for BrdU⁺DCX⁺ neurons.

(J) GSEA Heatmaps of core-enriched pathways; HD, PD, AD, and prion disease pathways were the top four neurologic disease pathways without or with rIFN- β , $n = 3$.

(K and L) IHC TH staining of coronal (K) STR and (L) SN sections. Data are (K) mean optical density (OD) \pm SEM and (L) TH⁺ cell bodies per SN area \pm SEM, $n = 4$ mice.

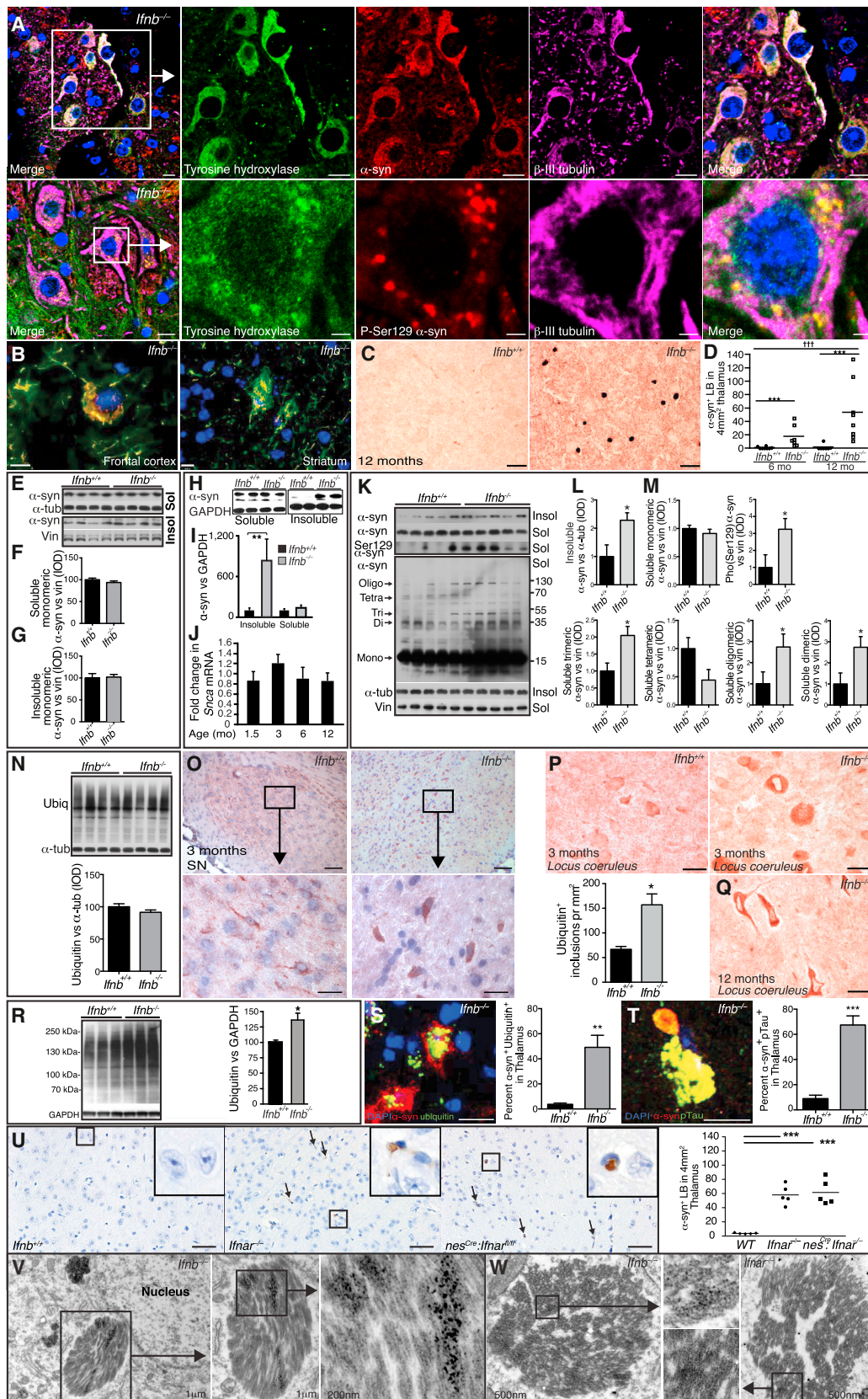
(M and N) Stitched IF images showing NeuN (green) and TH (red) (M) immunoreactivity and (N) quantifications of $n = 4$ mice/group in three to four ventral midbrain regions.

(O) WB of BG from 8-month-old mice ($n = 4$). Graph, mean integral optical density (IOD) \pm SEM of TH bands.

(P) Overview of the DA system. Red, upregulated; blue, downregulated genes from comparative microarray analysis of *Ifnb*^{+/-} and *Ifnb*^{-/-} CGNs.

(Q) High-performance liquid chromatography (HPLC) analysis of DA, DOPAC, and HVA in STR of 3- to 6-month-old mice, $n = 9$ –10/group.

For (I), (K), (L), (N), and (Q), $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, by Student's t test. See also Figure S2 and Data S1.



(legend on next page)

cerebellum, granular and molecular layers (Figure S3N), and significantly in thalamus of *lfnb*^{-/-} and *nes*^{Cre}:*lfnb*^{fl/fl} mice (Figure 3U). The ultrastructure of *lfnb*^{-/-} and *lfnar*^{-/-} neurons in paratenial and central medial thalamic nuclei contained α -syn-immunoreactive perinuclear LB-like structures (Figures 3V and 3W), which were absent in *lfnb*^{+/+} mice (data not shown), underscoring the importance of endogenous IFN- β -IFNAR signaling in preventing neuronal proteinopathy.

Lack of IFN- β Affects Autophagy

We used GSEA to identify cellular pathways involved in *lfnb*^{-/-} neuron pathology. In the top 20 deregulated pathways, three were associated with autophagy, which were restored with rIFN- β (Figure 4A; Table S2). The autophagy system uses adaptor proteins SQSTM1/p62 (hereafter p62) and NBR1 to bind ubiquitinated proteins and LC3B-II in the autophagosomal isolation membrane. Upon autophagy, cytosolic LC3B-I is converted to membrane bound LC3B-II and blocking autophagy flux, e.g., with Rab7 defects (Gutierrez et al., 2004; Hyttinen et al., 2013), causes accumulation of autophagy-targeted proteins; p62, NBR1, LC3B-II, and organelles. Increased neurons with accumulated p62 were seen in 3-month-old *lfnb*^{-/-} brains compared to WT, predominantly in brainstem, without increased p62 mRNA. LC3B-II increased in BG of 1.5-month-old *lfnb*^{-/-}, correlating with increased p62, NBR1 and Rab7, supporting defects in autophagy before α -syn, ubiquitin, pTau, and LB aggregation. Autophagy was even more deregulated in 6- to 8-month *lfnb*^{-/-} brains (Figures 4B–4H). Defective autophagy was also confirmed in *lfnar*^{-/-} mice (Figure S4).

Senescent or damaged mitochondria are degraded by mitophagy. Large, cytoplasmic, electron-dense aggregates associated with lipid droplets were exclusively found in *lfnb*^{-/-} and *lfnar*^{-/-} thalamic neurons, which correlated with significantly more mitochondria than in WT (Figures 4I–4M). Cultured *lfnb*^{-/-} neurons had significantly lower mitochondrial membrane potential (MMP) than *lfnb*^{+/+} neurons, indicating senescent or damaged

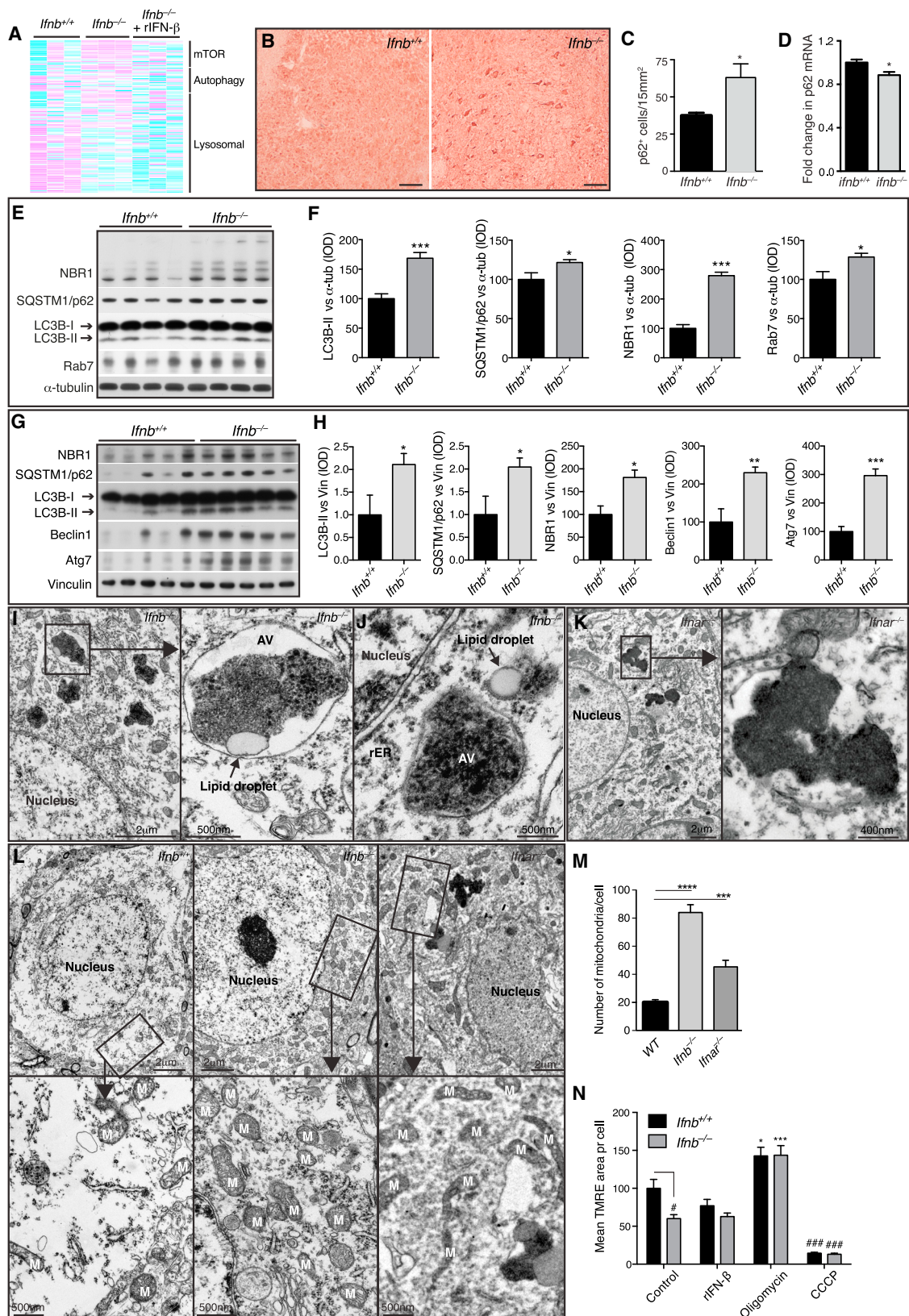
mitochondria. MMP was not affected by rIFN- β contrary to positive and negative regulators (Figure 4N). Thus, IFN- β signaling was important in regulating autophagy flux including neuronal mitophagy.

Late-Stage Autophagy Block Causes α -syn Accumulation

lfnb^{-/-} CNs and CGNs and *lfnar*^{-/-} CNs showed autophagy defects resembling a late-stage autophagy block (Figures 5A, 5B, S5A, and S5B). In *lfnb*^{-/-} CNs, LC3B-I and LC3B-II were increased with p62, NBR1, and K-63-linked ubiquitin, while K-48-linked ubiquitin only increased slightly (Figures 5A and 5B). Both *lfnb*^{+/+} and *lfnar*^{-/-} produced detectable IFN- β , unlike *lfnb*^{-/-} neurons (Figure S5C). To validate a block in autophagy flux, fusion between autophagosomes and lysosomes was inhibited with NH₄Cl. At 1 hr, LC3B-II levels saturated in both genotypes, which correlated with a significant increase in p62 in *lfnb*^{+/+} CNs; however, *lfnb*^{-/-} CNs had nearly saturated p62 levels without treatment, and no significant increase was seen with NH₄Cl (Figures 5C and 5D). Autophagy flux was measured in CNs with the mRFP-GFP-LC3B construct, which emits mRFP and GFP signal in autophagosomes and only mRFP in autolysosomes because of low pH in the latter. *lfnb*^{-/-} CNs had increased mCherry⁺GFP⁺ autophagosomes and very few autolysosomes compared to WT again suggesting blocked autophagy flux. Treating *lfnb*^{+/+} CNs with NH₄Cl caused a similar distribution as *lfnb*^{-/-} CNs. Rab7 significantly increased in *lfnb*^{-/-} CNs, suggesting accumulation of mature autophagosomes. In support, fewer LC3B- and p62-positive autophagic vacuoles overlapped with LAMP1, a lysosomal marker, in *lfnb*^{-/-} neurons indicating flux or fusion problems (Figures 5E–5H). GSEA suggested dysregulation in lysosomal genes in *lfnb*^{-/-} neurons (Table S2). However, numbers and dysfunction of lysosomes measured as acidification, morphology, or cathepsin expression and activity did not differ in *lfnb*^{+/+} and *lfnb*^{-/-} CNs and brains (Figures S5D–S5I).

Figure 3. *lfnb*^{-/-} Mice Develop LBs with α -syn, Ubiquitin, and pTau

- (A) IF of α -syn in SN of 6-month-old *lfnb*^{-/-} mice. Scale bar, 10 μ m; 2 μ m in close-up lower panel.
 (B) IF of 3-month-old *lfnb*^{-/-} brains showing α -syn⁺ neuron in frontal cortex and positive aggregates in STR. Scale bars, 10 μ m.
 (C) IHC of α -syn in thalamus. Scale bar, 20 μ m.
 (D) Number of α -syn-positive LBs in thalamus of 6- and 12-month-old mice. ***p < 0.001 by Mann-Whitney U test for two groups, †††p < 0.001 by Kruskal-Wallis test for all groups.
 (E–G) (E) WB of α -syn and quantified IOD in (F) soluble and (G) insoluble BG fractions of 1.5-month-old mice. Data are mean \pm SEM, n = 4.
 (H and I) (H) WB of α -syn and (I) quantified IOD in TX-100 soluble and insoluble fractions of 3-month-old brains. Data are mean \pm SEM, n = 5.
 (J) RT-PCR of α -syn in *lfnb*^{-/-} relative to *lfnb*^{+/+} brains. Data are mean \pm SEM, n = 3–5.
 (K) WB of α -syn (TX-100 soluble and insoluble) from BG of 8-month-old mice and long exposure of whole membranes of soluble fractions.
 (L and M) Quantified IOD of α -syn and pSer129- α -syn in (L) insoluble and (M) soluble fractions of monomeric (short exposure) and high-molecular-weight α -syn species (long exposure). Graphs, mean \pm SEM, n = 5.
 (N) WB and quantified ubiquitin IOD in BG from 1.5-month-old mice. Data are mean \pm SEM, n = 3.
 (O) IHC of ubiquitin in SN at 3 months and quantified ubiquitin⁺ aggregates/ μ m². Scale bar, 200 μ m; 20 μ m in inserts. Graph, mean \pm SEM, n = 5.
 (P and Q) IHC of ubiquitin in locus coeruleus at (P) 3 months and (Q) 12 months. Scale bar, 200 μ m.
 (R) WB and quantified ubiquitin IOD in brain extracts from 3-month-old mice. Data are mean \pm SEM, n = 3.
 (S and T) IF and quantification of (S) ubiquitin⁺ α -syn⁺ and (T) p-tau⁺ α -syn⁺ aggregates in thalami of 12-month-old mice. Scale bar, 10 μ m; data are percentage double positives \pm SEM.
 (U) IHC of α -syn in mice thalami showing LB-like structures (arrows). Scale bar, 50 μ m. Graph, α -syn⁺ LB-like structures in thalamus \pm SEM, n = 5. ***p < 0.001 by one-way ANOVA and Turkey's post hoc correction test.
 (V and W) Immuno-EM of thalamic neurons from *lfnb*^{-/-} and *lfnar*^{-/-} mice showing LB-like structures positive for α -syn by (V) immuno-DAB reactivity and (W) immunogold labeling.
 For (I), (L), (M), (O), (R), (S), and (T), *p < 0.05, **p < 0.01, ***p < 0.001 by Student's t test. See also Figure S3.



(legend on next page)

lfnb^{-/-} CNs developed ubiquitin⁺ α -syn aggregates and increased monomeric and high-molecular-weight α -syn after 21 days of culture (Figures 6A and 6B). Accumulation of α -syn was not due to proteasomal defects; turnover of the proteasome substrate p53 after cycloheximide treatment and proteasomal catalytic activity were uncompromised in *lfnb*^{-/-} CNs and unchanged by rIFN- β (Figure S6).

Blocking autophagy flux increased LC3B-II and α -syn, and α -syn colocalized with LC3B⁺ and p62⁺ autophagosomes in *lfnb*^{-/-} CNs (Figures 6C and 6D), underscoring the relevance for autophagic clearance of α -syn. Aggregation-prone pSer129- α -syn was seen in LC3B⁺ autophagosomes, that rarely overlapped with LAMP1. Collectively, these results demonstrated that lack of IFN- β reduced lysosomal fusion and caused α -syn accumulation.

IFN- β Promotes Neuronal Autophagy and α -syn Clearance

LC3B-II and p62 were higher in untreated *lfnb*^{-/-} CNs, and, while overnight rIFN- β -treatment promoted LC3B-II conversion and reduced p62 in *lfnb*^{+/-} CNs, indicating increased autophagy flux, rIFN- β reduced p62 but only slightly increased LC3B-II in *lfnb*^{-/-} CNs. By promoting autophagy, rIFN- β reduced α -syn in both *lfnb*^{+/-} and *lfnb*^{-/-} CNs (Figures 7A–7D). Effects of rapamycin, an mTOR-dependent autophagy activator, were similar to rIFN- β , but the mTOR-independent inducer trehalose more efficiently reduced α -syn in CNs from both genotypes, possibly through increased α -syn secretion (Ejlerskov et al., 2013). To ensure reconstitution of the genetic defect, we cultured *lfnb*^{-/-} neurons with a low rIFN- β dose for 21 days, which increased autophagy flux and α -syn clearance to the level in *lfnb*^{+/-} neurons (Figures 7E and 7F).

lfnb Gene Therapy Prevents Dopaminergic Neuron Loss in a Familial PD Model

We used lentiviruses to overexpress IFN- β to examine effects on a familial PD model induced with human α -syn (hSCNA) in rat SNs. Injection of hSCNA and control lentiviruses blocked autophagy indicated by accumulated LC3B-II, p62, NBR1, Beclin1, and hSCNA in BG 10 days after SN injection (Decressac et al., 2013) (Figures 7G and 7H). *lfnb* overexpression prevented hSCNA and pSer129- α -syn accumulation and restored TH loss. The mice showed improved left paw use compared to right paw use; contralateral to injection side of hSCNA/*lfnb* and

hSCNA/control viruses, respectively, 21 days post injection, which was associated with preservation of TH⁺ fibers in SN (Figures 7I and 7J). *lfnb* gene therapy also significantly protected TH⁺ dopaminergic neurons from hSCNA-induced SN damage (Figure 7K). Thus, IFN- β prevented pathology in a familial model of PD by inducing autophagy and α -syn clearance.

DISCUSSION

CNS immune activation and inflammation occur in neurodegenerative diseases (Brochard et al., 2009; Lee et al., 2009; Maccioni et al., 2009; Tansey et al., 2008), but their role in initiation is unclear. We report that defects in IFN- β -IFNAR signaling, that is central to immune regulation, trigger neurodegeneration in the CNS of aging mice. IFN- β promotes neurite growth and protein degradation by autophagy. Lack of IFN- β causes neural pathological changes: accumulation of LB-like structures, neural apoptosis, and neurogenesis defects. p62 accumulated in *lfnb*^{-/-} brainstems and BG regions before pathological α -syn aggregation. Aggregates and LB-like structures with α -syn were seen in brainstem and BG including SN, cortex, thalamus, and cerebellum as previously reported in PD and DLB patients (Goedert et al., 2013; Mori et al., 2003).

IFN- β is expressed by neurons to prevent malignant growth (Liu et al., 2013) and neuroinflammation (Liu et al., 2014) and is produced by choroid plexus epithelial cells in aged mouse and human brains without CNS disease (Baruch et al., 2014). In the latter study, injection of an anti-IFNAR antibody in the cerebrospinal fluid, however, positively affected some aspects of cognition. This approach could have complications as injection of full-length antibody might initiate complement-dependent cytotoxicity (Nelson, 2010) and inflammation (Congdon et al., 2013; Linnartz et al., 2012) and cause rapid antibody clearance compared to isotype controls (Sheehan et al., 2006). Genetic generation of *lfnb* or *lfnar* knockout mice or specifically targeting neuroectodermal cells in *nes*^{Cre}:*lfnar*^{f/f} mice circumvents such issues. We found that neuronal IFN- β -IFNAR signaling is required for neurons to withstand age-associated pathology. The data also suggest that neuronal IFN- β production in the CNS parenchyma might function differently than high production by choroid plexus epithelial cells or potentially resident microglia.

Familial but not sporadic neurodegenerative diseases are associated with overexpressed or mutant proteins such as

Figure 4. Late-Stage Degradative Autophagosomes Accumulate in *lfnb*^{-/-} Neurons

(A) Heatmaps of GSEA core-enriched genes in autophagy-related pathways comparing *lfnb*^{+/-} and *lfnb*^{-/-} mouse CGNs, *lfnb*^{-/-} with or without 24 hr rIFN- β (100 U/ml), n = 3.

(B) IHC from pons area using p62 antibodies. Scale bar, 20 μ m.

(C) Quantification of p62⁺ cells/15 mm². Data are mean \pm SEM, n = 5.

(D) Fold change in p62 mRNA in brain extracts of 3-month-old mice. Data are mean \pm SEM, n = 3.

(E–H) WB of BG of (E) 1.5-month-old and (G) 6- to 8-month-old mice and (F) and (H) quantified IOD of bands. Data are mean \pm SEM of four or five brains.

(I–L) TEM of 9-month-old mice thalami. (I–J) *lfnb*^{-/-} mice showing perinuclear electron-dense late-stage autophagic vacuoles (AV), most surrounded by single lipid membrane and associated with lipid droplets. rER, rough endoplasmic reticulum. (K) *lfnar*^{-/-} mouse with similar electron-dense aggregates. (L) TEM of mitochondria in thalamic neuron cell body. M, mitochondria.

(M) Graph, mean \pm SEM mitochondria per cell body of 12–16 thalamic neurons; ***p < 0.001, ****p < 0.0001 by one-way ANOVA with post hoc Dunnett's multiple comparisons test.

(N) TMRE analysis of 21-day-old CNs with mean TMRE area per cell, n = 3. *Treatment effect, and #genotype effect.

For (B), (C), (F), (H), and (N), */#p < 0.05, **p < 0.01, ***/#p < 0.001 by unpaired Student's t test. See also Figure S4 and Table S2.

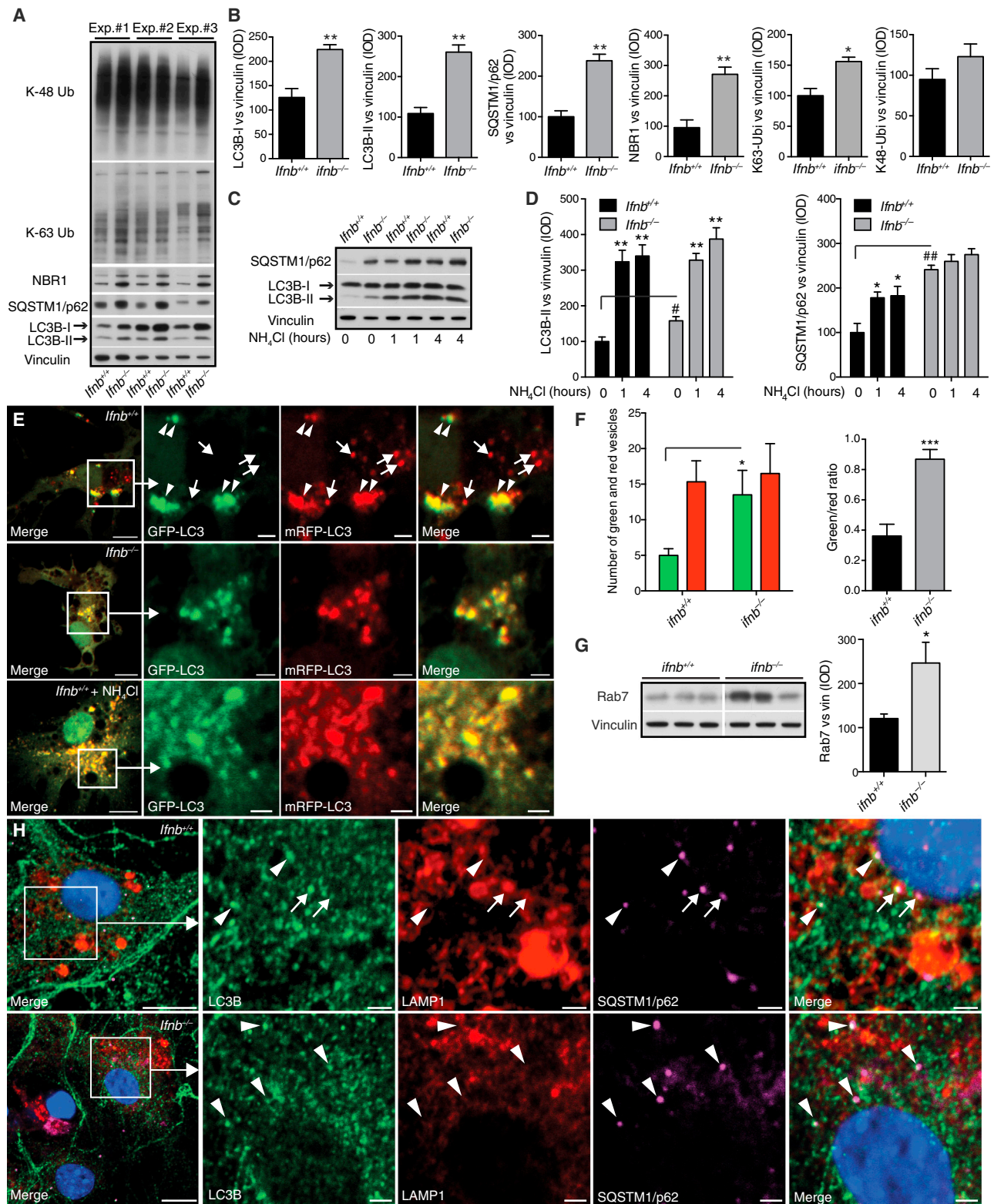


Figure 5. *Ifnb*^{-/-} Neurons Have a Defect in Autophagy Maturation

(A–H) *Ifnb*^{+/+} and *Ifnb*^{-/-} primary CNs cultured for 21 days.

(A) WB from three independent experiments with antibodies as indicated.

(legend continued on next page)

α -syn in PD (Thenganatt and Jankovic, 2014). Overexpression of disease-associated proteins (Arima et al., 1999; Jellinger, 2000; Lipka et al., 1998) was not required to accumulate LB-like structures in *lfnb*^{-/-}, *lfnar*^{-/-}, or *nes*^{Cre}:*lfnar*^{fl/fl} neurons. *lfnb* or *lfnar* deletion was the sole trigger for α -syn-containing inclusion bodies resulting from defective autophagosome maturation. Treatment with rIFN- β promoted autophagy flux and increased α -syn clearance in *lfnb*^{+/+} and *lfnb*^{-/-} neurons, supporting IFN- β function in mutated proteins clearance, as suggested with ataxin 7 (Chort et al., 2013). Shared pathology of *lfnb*^{-/-} and *nes*^{Cre}:*lfnar*^{fl/fl} mice suggests that the pathogenesis is likely driven by initial changes in neuroectodermal neurons rather than systemic or local immune activation. In agreement, neurons are crucial for regulation of CNS inflammation (Liu et al., 2006).

Protein degradation defects are common in neurodegenerative pathologies (Davies et al., 1997; Hara et al., 2006; Komatsu et al., 2006). We found that endogenous neuronal IFN- β signaling is central in regulating protein degradation by autophagy, including clearing aged mitochondria. Accumulated aged and defective mitochondria may release reactive oxygen species, enhancing neuroinflammation and neuronal death (Chaturvedi and Flint Beal, 2013). Cognitive and motoric impairments in PD is associated with DA dysfunction (Narayanan et al., 2013). Excess cytosolic DA is degraded by monoamine oxidase (MAO) in the mitochondrial outer membrane and MAO defects might cause oxidative stress (Segura-Aguilar et al., 2014). Senescent and damaged mitochondria and reduced MAO mRNA were found in *lfnb*^{-/-} mice and reduced DOPAC in *lfnb*^{-/-} mice supported dysregulation of MAO and mitochondrial genes potentially contributing to neurotoxicity and DA neuron death.

Lack of endogenous IFN- β signaling was associated with spontaneous neurodegeneration, impaired motor coordination and cognition, and neuronal LB-like inclusions with aging as seen in most PD and DLB patients (Jellinger, 2008). Gene profiling supported essential IFN- β regulation of neuronal homeostasis. Neurotrophin was a top hit in GSEA and increased Tox mRNA, important for corticogenesis (Artegiani et al., 2015), supported IFN- β effects on prosurvival mechanisms.

We showed that *lfnb* gene therapy reversed pathology in a familial PD model, by promoting autophagy and α -syn clearance, which preserved DA neurons and associated neurologic deficit. Our data strongly support an essential role for IFN- β signaling in preventing neurodegenerative pathology and suggest *lfnb*^{-/-} mice as a model for nonfamilial, sporadic neurodegenerative dis-

eases, particularly PD and DLB, with potential for testing future therapies.

EXPERIMENTAL PROCEDURES

For detailed procedures, see Supplemental Experimental Procedures.

Mice and Cell Culture

lfnb^{-/-} mice (Erlandsson et al., 1998) were backcrossed 20 generations to B10.RIII or C57BL6. *lfnar*^{-/-} and *nes*^{Cre}:*lfnar*^{fl/fl} mice were in C57BL6 (Prinz et al., 2008). WT were *lfnb*^{+/-}, *lfnb*^{+/+} littermates, or *lfnar*^{+/+} C57BL6 mice. C57BL/6JOLA-Hsd mice (Harlan Laboratories) with a spontaneous deletion of part of *Snca* (α -syn) were crossed with *lfnb*^{-/-} mice for *Snca*^{-/-}*lfnb*^{-/-}. Mice were housed in standard facilities. Sex- and weight-matched mice were used in experiments performed in accordance with the ethical committees in Denmark and approved by our institutional review boards. CGNs were from 6- or 7-day-old cerebella and cortical neuron (CN) cultures from the cortex of 1-day-old mice.

Behavioral Measurements

Motor-coordination and -learning were evaluated with an accelerating RotaRod (TSE Systems GmbH) automatically recording time before fall. Neuromuscular strength was tested by forelimb hanging time on a bar. Heat and cold tail-pain sensitivity was measured by tail-flick latency time after exposure.

Spatial learning and reference memory were assessed with Morris water maze (Vorhees and Williams, 2006) with slight modifications. Swimming patterns were recorded with Ethovision 3.1 (Noldus Information Technology), measuring the time to reach a hidden platform during learning trials and frequency of platform position crossings during probe tests. Deficits prevalence were calculated as behavioral test scores of *lfnb*^{-/-} mice that deviated from the norm: mean value (SD/2) of the *lfnb*^{+/+} group (Table S1).

In the cylinder tests, asymmetry in forelimb use during vertical exploration was used as a validated measure of akinesia in hemiparkinsonian rodents.

Cloning

Mouse *lfnb* pCR4IFNb was from transOMIC (accession no. BC119395). *lfnb* was transferred to pCSII-GW via pCR8TOPOGW (Invitrogen) with conventional cloning techniques to generate pCSII-IFNb. PCSII (without insert) was generated by lambda recombination with an empty pCR8GW vector. Plasmid inserts were verified by sequencing.

Surgery with AAV6-hSCNA

AAV-GFP or AAV-human α -syn/hSNCA-WPRE (Vector Biolabs) together with lentiviral vectors pCSII-IFNb and pCSII control (3 μ l for each virus) were injected unilaterally into 30 adult female Sprague Dawley rats (Taconic; 225–250 g at surgery). Virus was infused at 0.2 μ l/min as described (Decressac et al., 2013).

Immunohistochemistry, Immunofluorescence, and Transmission Electron Microscopy

For immunohistochemistry (IHC) and immunofluorescence (IF), either mice were perfused and brains fixed in 4% paraformaldehyde (PFA) and paraffin

(B) Quantified IOD of WB bands normalized to vinculin. Data are mean \pm SEM, $n = 3$.

(C) WB of CNs with or without NH₄Cl (20 mM) for 1 or 4 hr.

(D) Quantified IOD of WB. Data are mean \pm SEM of $n = 3$, * $p < 0.05$, ** $p < 0.01$ by one-way ANOVA. *Within and # between genotype differences after NH₄Cl treatment.

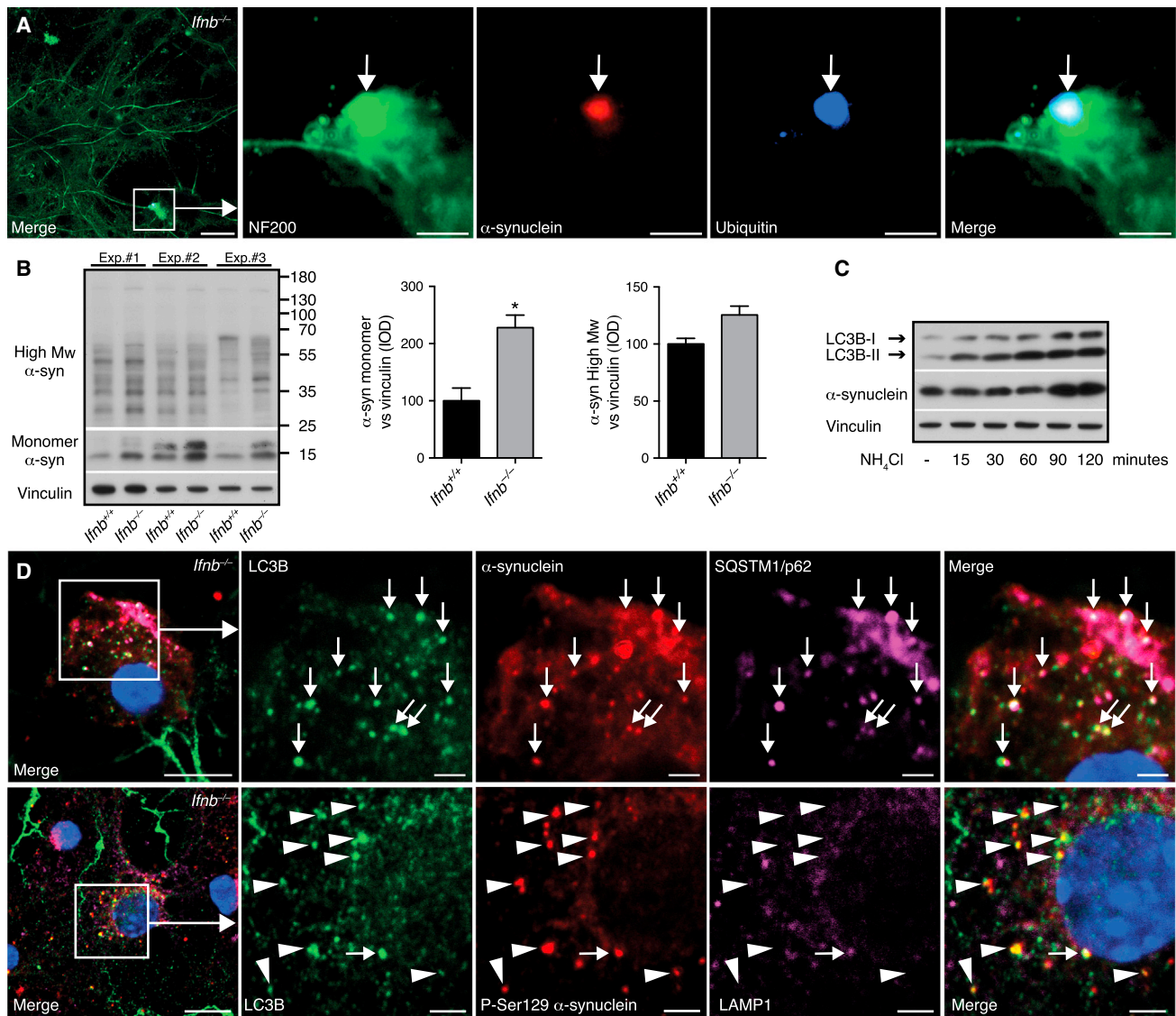
(E) CNs expressing mRFP-GFP-LC3B with or without NH₄Cl for 2 hr. Arrowheads, colocalized GFP and mRFP (autophagosomes); arrows, mRFP-only vesicles (autolysosomes). Scale bars, 10 μ m; 2 μ m in inserts.

(F) Graphs, mean number of autophagosomes and autolysosomes/cell and vesicle ratio \pm SEM, $n = 3$.

(G) WB of Rab7. Graph, IOD of WB. Data are mean \pm SEM of $n = 3$.

(H) IF showing LC3B and p62 colocalized in autophagosomes (arrowheads) and LC3B, LAMP1, and p62 colocalizing in autolysosomes (arrows). Scale bars, 10 μ m; 2 μ m in inserts.

For (B), (F), and (G), * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$ by unpaired Student's t test. See also Figure S5.



embedded or brains were dissected and snap-frozen before sectioning. In vitro neuronal cultures were fixed in 4% PFA before staining. Tissues and cells were stained as described (Liu et al., 2014).

For neurogenesis and BrdU staining, mice were injected intraperitoneally (i.p.) once per day (75 μ g/g body weight) for 5 consecutive days and sacrificed 2 hr after last injection. Brains were processed and immunostained for BrdU and doublecortin (DCX).

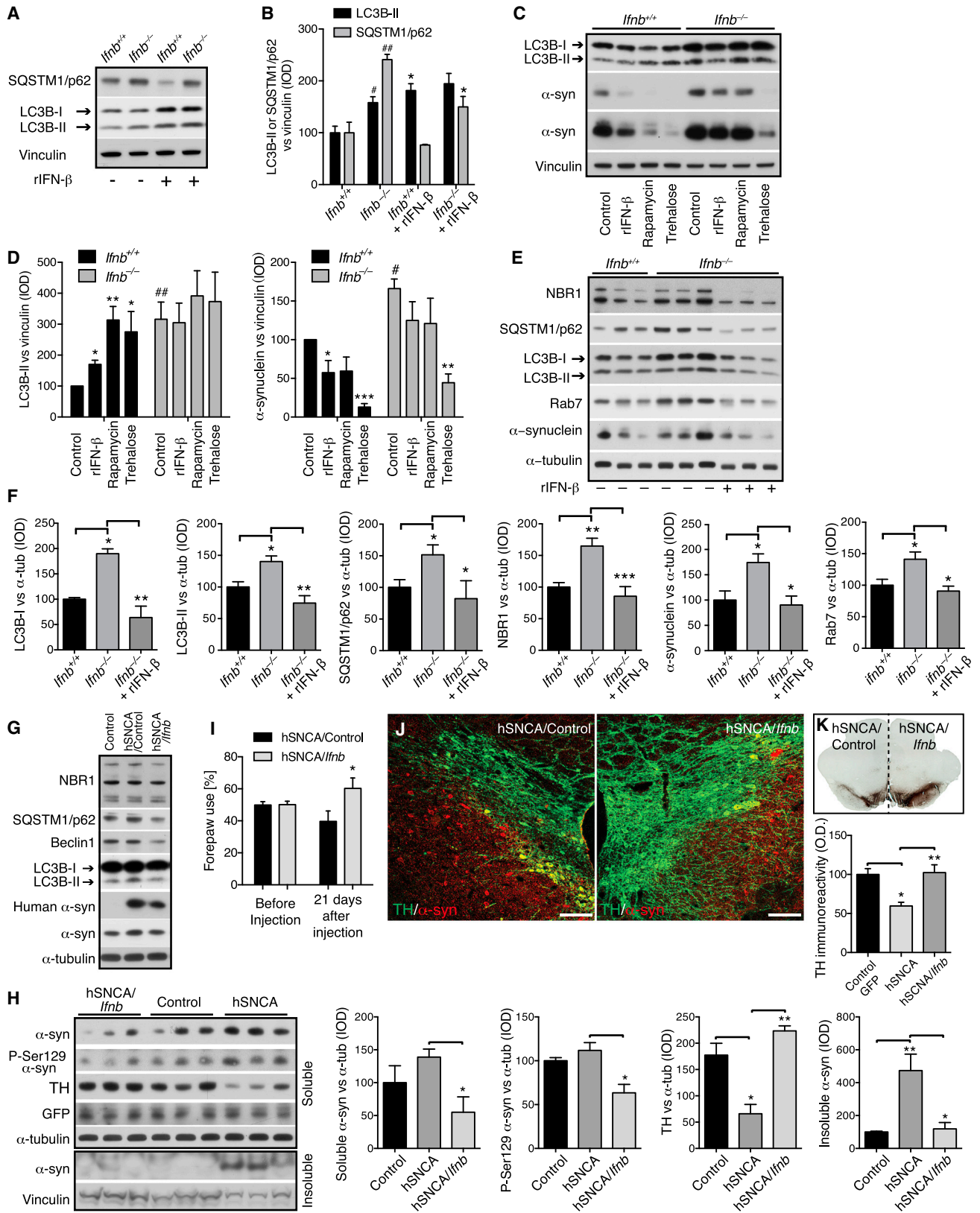
MMP was measured by adding tetramethylrhodamine ethyl ester (TMRE) and Hoechst (Life Technologies) to primary CNs.

IF images were taken with a Zeiss LSM510 confocal scanning microscope and IN Cell Analyzer 2200 automated microscope. IHC images were taken with a NanoZoomer 2.0-HT digital slide scanner or Olympus BX51 micro-

scope. Images were quantified with ImageJ (Fiji version), IN Cell Investigator, CellProfiler, Zeiss Zen, and Adobe Photoshop.

In situ apoptosis detection was with TUNEL kits (Calbiochem) with Hoechst counterstaining or DAB substrate and methyl green counterstaining kits (R&D Systems).

For transmission electron microscopy (TEM), 12-month-old mice were cardiac perfused with 2% PFA and 0.2% glutaraldehyde and paratenial, and central medial thalamic nuclei were dissected and processed for epon embedding and ultrathin sectioning. Samples were incubated with primary α -syn antibodies (Leica) and biotinylated (Vector Laboratories) or 1.4-nm gold-labeled secondary antibodies (Nanoprobes) and embedded in epoxy resin. Ultra-thin sections were analyzed in a Philips CM100 electron microscope.



(legend on next page)

Transfection and Plasmids

CNs were transfected with mRFP-GFP-LC3 using Lipofectamine 2000 (Life Technologies) according to the manufacturer's description for live-cell confocal imaging.

Stereological Analysis

The optical fractionator method was used to estimate neuron and glial numbers in the hippocampus of 80- μ m mouse brain sections stained with H&E. Pointcounting techniques, based on the Cavalieri principle, were used to estimate hippocampal volume on one side of the brain.

Real-Time PCR

Total RNA was isolated using a QIAGEN kit (QIAGEN), reverse transcribed into cDNA, amplified, and quantified by SYBR Green (Bio-Rad) detection. Relative mRNA expression was normalized with glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) gene.

High-Performance Liquid Chromatography

Striatum was dissected and homogenized. Filtered supernatant was examined for DA, DOPAC, and HVA levels by reversed-phase HPLC (Decressac et al., 2013).

Western Blots

Samples were lysed in 1% Triton X-100 (Sigma) and insoluble brain pellets were sonicated in UREA/SDS and processed as described (Ejlertsen et al., 2013).

Affymetrix Microarrays

RNA was extracted with TRI (Sigma) and DNase I (Invitrogen) from 3-day-old *Ifnb*^{+/+} and *Ifnb*^{-/-} CGN cultures in triplicate with or without 24 hr rIFN- β (100 U/ml). Affymetrix 430 2.0 microarray chip (SCIBLU, Affymetrix) data were analyzed with Arraystar 3 (DNA STAR) and quantile-normalized and processed by the RMA (Affymetrix) algorithm. We log₂-transformed intensity values, and normal-distributed data were tested in unpaired two-tailed Student's t tests, filtering for differential regulation confidence of 95% ($p < 0.05$). Venn diagrams were created with oneChannelGUI (Bioconductor). Quantile-normalized RMA-treated data selected using a 1.4-fold cutoff were analyzed with Ingenuity Pathway Analysis software.

GSEA data heatmaps were generated by extracting lists of core enriched genes from GSEA pathway analysis in R using Heatmap2 (Data S1).

Comparisons of *Ifnb*^{-/-} CGN gene profiles and published PD and HD models using Affymetrix data (GSE4758, GSE9038) were quantile-normalized together and summarized in R using RMA algorithms from the Affy-package. Differential expression was determined individually within each experiment for our data and published PD and HD models by comparing control/WT samples with transgenic samples using standard ANOVA.

Statistical Analysis

Data were analyzed with unpaired and paired two-tailed Student's t tests, ANOVA, and Mann-Whitney U, and Kruskal-Wallis tests. $p < 0.05$ was significant. Error bars are SEM.

ACCESSION NUMBERS

The accession number for the raw microarray data reported in this paper is Gene Expression Omnibus: GSE63815.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, two tables, and a data file and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.069>.

AUTHOR CONTRIBUTIONS

P.E., J.G.H., J.W., R.C., M.K., Y.L., G.P., K.K., M.A., C.F.R., and K.R. did experiments and analyzed and prepared data; B.P. contributed to stereological studies; T.G., D.L., and M.P. did experiments including immunolabeling-EM and *Ifnar*^{-/-} and *nes*^{Cre}:*Ifnar*^{fl/fl} mouse analysis; D.C.R. contributed material and designed some protein degradation experiments and S.I.-N. designed and supervised the study, analyzed and interpreted data, and wrote the manuscript. All authors read and contributed to the final manuscript.

ACKNOWLEDGMENTS

Support to S.I.-N. was from Danish Council For Independent Research (DFF)-Medical Sciences, Alzheimer-forskningsfonden, Danish Multiple Sclerosis Society, Danish Cancer Society, and Lundbeck Foundation. D.C.R. is a Wellcome Trust Principal Research Fellow. The authors thank Marja Jäättelä for help establishing lysosomal assays, Janne Kock for efficient experimental animal facility arrangement, Elin Pietras for help with IN Cell analysis, Mohamoud Osman, Alexandra Müller, Ana Rita Antunes, Sigrun Nestel, and Barbara Joch for technical assistance, and Klaus Qvortrup for electron microscopy help.

Received: December 4, 2014

Revised: May 20, 2015

Accepted: August 13, 2015

Published: October 8, 2015

REFERENCES

Arima, K., Hirai, S., Sunohara, N., Aoto, K., Izumiyama, Y., Uéda, K., Ikeda, K., and Kawai, M. (1999). Cellular co-localization of phosphorylated tau- and

Figure 7. IFN- β Treatment Promotes Autophagy Flux and Reduces α -syn Accumulation

(A–F) *Ifnb*^{+/+} and *Ifnb*^{-/-} primary CNs cultured for 21 days.

(A) WB of CN with or without rIFN- β (100 U/ml) for 22 hr.

(B) Quantified IOD of WB bands. Graphs, mean \pm SEM, $n = 3$; * $p < 0.05$; ** $p < 0.01$ by one-way ANOVA. #Between genotypes and *after rIFN- β in each group.

(C) WB of LC3, α -syn (two exposures), and vinculin from untreated or rIFN- β (100 U/ml), rapamycin (1 μ M), or trehalose (50 mM) -treated CNs (24 hr).

(D) Quantified IOD of WB bands with mean \pm SEM, $n = 4$ –5. *After treatment within genotype and #between untreated control genotypes.

(E) WB of CNs with 30 U/ml rIFN- β (every 3–4 days) and 100 U/ml the last 24 hr.

(F) Quantified IOD. Data are mean \pm SEM, $n = 4$ –6.

(G and H) WB of rat BG 10 days post AAV- and lentiviral-control or AAV-hSNCA co-injection with lentivirus-*Ifnb* or lentivirus-control vector. Antibodies were against autophagy markers, human α -syn, or human and rat α -syn, pSer129- α -syn, TH, and GFP (to confirm AAV and lentiviral expression). Graphs, mean \pm SEM, $n = 3$ /group.

(I) Right and left forepaw use in rats before and 21 days after AAV injection of hSNCA with lentivirus-control or lentivirus-*Ifnb* in left and right SN brain hemisphere, respectively. Graph, mean \pm SD, $n = 3$ –5 rats/group. * $p < 0.05$ by Student's t test.

(J) IF of rat brain; right and left hemisphere 21 days after virus injection. Scale bar, 100 μ m.

(K) IHC of rat 21 days after virus injection as in (I) with SN TH immunoreactivity. Graph, mean (OD) \pm SEM. TH immunoreactivity of control AAV/lentivirus, AAV-hSNCA/lentivirus control, or AAV hSNCA/lentivirus-*Ifnb*, $n = 3$ –5 rats/group.

For (B), (D), (F), (H), and (K), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by one-way ANOVA.

- NACP/alpha-synuclein-epitopes in lewy bodies in sporadic Parkinson's disease and in dementia with Lewy bodies. *Brain Res.* 843, 53–61.
- Artegianni, B., de Jesus Domingues, A.M., Bragado Alonso, S., Brandl, E., Mas-salini, S., Dahl, A., and Calegari, F. (2015). Tox: a multifunctional transcription factor and novel regulator of mammalian corticogenesis. *EMBO J.* 34, 896–910.
- Bartels, T., Choi, J.G., and Selkoe, D.J. (2011). α -Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* 477, 107–110.
- Baruch, K., Deczkowska, A., David, E., Castellano, J.M., Miller, O., Kertser, A., Berkutski, T., Barnett-Itzhaki, Z., Bezalel, D., Wyss-Coray, T., et al. (2014). Aging. Aging-induced type I interferon response at the choroid plexus negatively affects brain function. *Science* 346, 89–93.
- Brochard, V., Combadière, B., Prigent, A., Laouar, Y., Perrin, A., Beray-Berthat, V., Bonduelle, O., Alvarez-Fischer, D., Callebert, J., Launay, J.M., et al. (2009). Infiltration of CD4+ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J. Clin. Invest.* 119, 182–192.
- Chaturvedi, R.K., and Flint Beal, M. (2013). Mitochondrial diseases of the brain. *Free Radic. Biol. Med.* 63, 1–29.
- Chort, A., Alves, S., Marinello, M., Dufresnois, B., Dornbier, J.G., Tesson, C., Latouche, M., Baker, D.P., Barkats, M., El Hachimi, K.H., et al. (2013). Interferon β induces clearance of mutant ataxin 7 and improves locomotion in SCA7 knock-in mice. *Brain* 136, 1732–1745.
- Congdon, E.E., Gu, J., Sait, H.B., and Sigurdsson, E.M. (2013). Antibody uptake into neurons occurs primarily via clathrin-dependent Fc γ receptor endocytosis and is a prerequisite for acute tau protein clearance. *J. Biol. Chem.* 288, 35452–35465.
- Davies, S.W., Turmaine, M., Cozens, B.A., DiFiglia, M., Sharp, A.H., Ross, C.A., Scherzinger, E., Wanker, E.E., Mangiarini, L., and Bates, G.P. (1997). Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90, 537–548.
- Decressac, M., Mattsson, B., Weikop, P., Lundblad, M., Jakobsson, J., and Björklund, A. (2013). TFEB-mediated autophagy rescues midbrain dopamine neurons from α -synuclein toxicity. *Proc. Natl. Acad. Sci. USA* 110, E1817–E1826.
- Ejlerskov, P., Rasmussen, I., Nielsen, T.T., Bergström, A.L., Tohyama, Y., Jensen, P.H., and Vilhardt, F. (2013). Tubulin polymerization-promoting protein (TPPP/p25 α) promotes unconventional secretion of α -synuclein through exophagy by impairing autophagosome-lysosome fusion. *J. Biol. Chem.* 288, 17313–17335.
- Erlundsson, L., Blumenthal, R., Eloranta, M.L., Engel, H., Alm, G., Weiss, S., and Leanderson, T. (1998). Interferon-beta is required for interferon-alpha production in mouse fibroblasts. *Curr. Biol.* 8, 223–226.
- Fossale, E., Seong, I.S., Coser, K.R., Shioda, T., Kohane, I.S., Wheeler, V.C., Gusella, J.F., MacDonald, M.E., and Lee, J.M. (2011). Differential effects of the Huntington's disease CAG mutation in striatum and cerebellum are quantitative not qualitative. *Hum. Mol. Genet.* 20, 4258–4267.
- Francis, P.T. (2009). Biochemical and pathological correlates of cognitive and behavioural change in DLB/PDD. *J. Neurol.* 256 (Suppl 3), 280–285.
- Goedert, M., Spillantini, M.G., Del Tredici, K., and Braak, H. (2013). 100 years of Lewy pathology. *Nat. Rev. Neurol.* 9, 13–24.
- Gutierrez, M.G., Munafó, D.B., Berón, W., and Colombo, M.I. (2004). Rab7 is required for the normal progression of the autophagic pathway in mammalian cells. *J. Cell Sci.* 117, 2687–2697.
- Hara, T., Nakamura, K., Matsui, M., Yamamoto, A., Nakahara, Y., Suzuki-Migishima, R., Yokoyama, M., Mishima, K., Saito, I., Okano, H., and Mizushima, N. (2006). Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* 441, 885–889.
- Harris, H., and Rubinstein, D.C. (2012). Control of autophagy as a therapy for neurodegenerative disease. *Nat. Rev. Neurol.* 8, 108–117.
- Hyttinen, J.M., Niittykoski, M., Salminen, A., and Kaamiranta, K. (2013). Maturation of autophagosomes and endosomes: a key role for Rab7. *Biochim. Biophys. Acta* 1833, 503–510.
- Jellinger, K.A. (2000). Morphological substrates of mental dysfunction in Lewy body disease: an update. *J. Neural Transm. Suppl.* 59, 185–212.
- Jellinger, K.A. (2008). Neuropathological aspects of Alzheimer disease, Parkinson disease and frontotemporal dementia. *Neurodegener. Dis.* 5, 118–121.
- Jellinger, K.A., and Attems, J. (2008). Prevalence and impact of vascular and Alzheimer pathologies in Lewy body disease. *Acta Neuropathol.* 115, 427–436.
- Komatsu, M., Waguri, S., Chiba, T., Murata, S., Iwata, J., Tanida, I., Ueno, T., Koike, M., Uchiyama, Y., Kominami, E., and Tanaka, K. (2006). Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* 441, 880–884.
- Lee, Y.K., Kwak, D.H., Oh, K.W., Nam, S.Y., Lee, B.J., Yun, Y.W., Kim, Y.B., Han, S.B., and Hong, J.T. (2009). CCR5 deficiency induces astrocyte activation, A β deposit and impaired memory function. *Neurobiol. Learn. Mem.* 92, 356–363.
- Linnartz, B., Kopatz, J., Tenner, A.J., and Neumann, H. (2012). Sialic acid on the neuronal glycocalyx prevents complement C1 binding and complement receptor-3-mediated removal by microglia. *J. Neurosci.* 32, 946–952.
- Lippa, C.F., Fujiwara, H., Mann, D.M., Giasson, B., Baba, M., Schmidt, M.L., Nee, L.E., O'Connell, B., Pollen, D.A., St George-Hyslop, P., et al. (1998). Lewy bodies contain altered alpha-synuclein in brains of many familial Alzheimer's disease patients with mutations in presenilin and amyloid precursor protein genes. *Am. J. Pathol.* 153, 1365–1370.
- Liu, Y., Teige, I., Birnir, B., and Issazadeh-Navikas, S. (2006). Neuron-mediated generation of regulatory T cells from encephalitogenic T cells suppresses EAE. *Nat. Med.* 12, 518–525.
- Liu, Y., Carlsson, R., Ambjörn, M., Hasan, M., Badn, W., Darabi, A., Siesjö, P., and Issazadeh-Navikas, S. (2013). PD-L1 expression by neurons nearby tumors indicates better prognosis in glioblastoma patients. *J. Neurosci.* 33, 14231–14245.
- Liu, Y., Carlsson, R., Comabella, M., Wang, J., Kosicki, M., Carrion, B., Hasan, M., Wu, X., Montalban, X., Dziegiel, M.H., et al. (2014). FoxA1 directs the lineage and immunosuppressive properties of a novel regulatory T cell population in EAE and MS. *Nat. Med.* 20, 272–282.
- Maccioni, R.B., Rojo, L.E., Fernández, J.A., and Kuljis, R.O. (2009). The role of neuroimmunomodulation in Alzheimer's disease. *Ann. N Y Acad. Sci.* 1153, 240–246.
- Mori, F., Piao, Y.S., Hayashi, S., Fujiwara, H., Hasegawa, M., Yoshimoto, M., Iwatsubo, T., Takahashi, H., and Wakabayashi, K. (2003). Alpha-synuclein accumulates in Purkinje cells in Lewy body disease but not in multiple system atrophy. *J. Neuropathol. Exp. Neurol.* 62, 812–819.
- Mrak, R.E., and Griffin, W.S. (2007). Dementia with Lewy bodies: Definition, diagnosis, and pathogenic relationship to Alzheimer's disease. *Neuropsychiatr. Dis. Treat.* 3, 619–625.
- Narayanan, N.S., Rodnitsky, R.L., and Uc, E.Y. (2013). Prefrontal dopamine signaling and cognitive symptoms of Parkinson's disease. *Rev. Neurosci.* 24, 267–278.
- Nelson, A.L. (2010). Antibody fragments: hope and hype. *MAbs* 2, 77–83.
- Prinz, M., Schmidt, H., Mildner, A., Knobloch, K.P., Hanisch, U.K., Raasch, J., Merkler, D., Detje, C., Gutcher, I., Mages, J., et al. (2008). Distinct and nonredundant in vivo functions of IFNAR on myeloid cells limit autoimmunity in the central nervous system. *Immunity* 28, 675–686.
- Rockenstein, E., Nuber, S., Overk, C.R., Ubhi, K., Mante, M., Patrick, C., Adame, A., Trejo-Morales, M., Gerez, J., Picotti, P., et al. (2014). Accumulation of oligomer-prone α -synuclein exacerbates synaptic and neuronal degeneration in vivo. *Brain* 137, 1496–1513.
- Segura-Aguilar, J., Paris, I., Muñoz, P., Ferrari, E., Zecca, L., and Zucca, F.A. (2014). Protective and toxic roles of dopamine in Parkinson's disease. *J. Neurochem.* 129, 898–915.
- Sheehan, K.C., Lai, K.S., Dunn, G.P., Bruce, A.T., Diamond, M.S., Heutel, J.D., Duno-Arthur, C., Carrero, J.A., White, J.M., Hertzog, P.J., and Schreiber, R.D.

(2006). Blocking monoclonal antibodies specific for mouse IFN- α /beta receptor subunit 1 (IFNAR-1) from mice immunized by in vivo hydrodynamic transfection. *J. Interferon Cytokine Res.* **26**, 804–819.

Shulman, J.M., and De Jager, P.L. (2009). Evidence for a common pathway linking neurodegenerative diseases. *Nat. Genet.* **41**, 1261–1262.

Tansey, M.G., Frank-Cannon, T.C., McCoy, M.K., Lee, J.K., Martinez, T.N., McAlpine, F.E., Ruhn, K.A., and Tran, T.A. (2008). Neuroinflammation in Parkinson's disease: is there sufficient evidence for mechanism-based interventional therapy? *Front. Biosci.* **13**, 709–717.

Teige, I., Treschow, A., Teige, A., Mattsson, R., Navikas, V., Leanderson, T., Holmdahl, R., and Issazadeh-Navikas, S. (2003). IFN- β gene deletion leads to augmented and chronic demyelinating experimental autoimmune encephalomyelitis. *J. Immunol.* **170**, 4776–4784.

Teige, I., Liu, Y., and Issazadeh-Navikas, S. (2006). IFN- β inhibits T cell activation capacity of central nervous system APCs. *J. Immunol.* **177**, 3542–3553.

Thenganatt, M.A., and Jankovic, J. (2014). Parkinson disease subtypes. *JAMA Neurol.* **71**, 499–504.

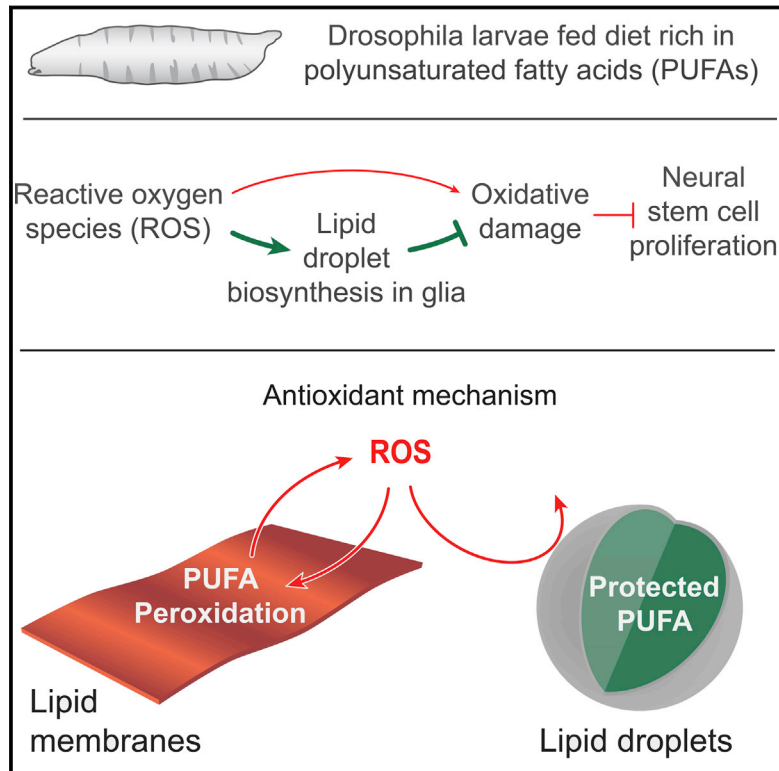
Vorhees, C.V., and Williams, M.T. (2006). Morris water maze: procedures for assessing spatial and related forms of learning and memory. *Nat. Protoc.* **1**, 848–858.

Wang, W., Perovic, I., Chittuluru, J., Kaganovich, A., Nguyen, L.T., Liao, J., Auclair, J.R., Johnson, D., Landeru, A., Simorellis, A.K., et al. (2011). A soluble α -synuclein construct forms a dynamic tetramer. *Proc. Natl. Acad. Sci. USA* **108**, 17797–17802.

Yong, V.W., Chabot, S., Stuve, O., and Williams, G. (1998). Interferon beta in the treatment of multiple sclerosis: mechanisms of action. *Neurology* **51**, 682–689.

Antioxidant Role for Lipid Droplets in a Stem Cell Niche of *Drosophila*

Graphical Abstract



Authors

Andrew P. Bailey, Grielof Koster, Christelle Guillermer, ..., Claude P. Lechene, Anthony D. Postle, Alex P. Gould

Correspondence

alex.gould@crick.ac.uk

In Brief

Lipid droplets can act as antioxidant organelles that protect *Drosophila* neural stem cells from hypoxia-triggered ROS. They sequester polyunsaturated fatty acids away from plasma membranes, thereby shielding these lipids from harmful peroxidation that would otherwise compromise proliferation.

Highlights

- Oxidative stress stimulates lipid droplet biosynthesis in a neural stem cell niche
- Lipid droplets protect niche and neural stem cells from damaging PUFA peroxidation
- PUFAs are less vulnerable to peroxidation in lipid droplets than in cell membranes



Antioxidant Role for Lipid Droplets in a Stem Cell Niche of *Drosophila*

Andrew P. Bailey,¹ Grielof Koster,² Christelle Guillermier,³ Elizabeth M.A. Hirst,¹ James I. MacRae,¹ Claude P. Lechene,³ Anthony D. Postle,² and Alex P. Gould^{1,*}

¹The Francis Crick Institute, Mill Hill Laboratory, The Ridgeway, Mill Hill, London NW7 1AA, UK

²Academic Unit of Clinical & Experimental Sciences, Faculty of Medicine, Sir Henry Wellcome Laboratories, Southampton General Hospital, Southampton SO16 6YD, UK

³National Resource for Imaging Mass Spectroscopy, Harvard Medical School and Brigham and Women's Hospital, Cambridge, MA 02139, USA

*Correspondence: alex.gould@crick.ac.uk

<http://dx.doi.org/10.1016/j.cell.2015.09.020>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

SUMMARY

Stem cells reside in specialized microenvironments known as niches. During *Drosophila* development, glial cells provide a niche that sustains the proliferation of neural stem cells (neuroblasts) during starvation. We now find that the glial cell niche also preserves neuroblast proliferation under conditions of hypoxia and oxidative stress. Lipid droplets that form in niche glia during oxidative stress limit the levels of reactive oxygen species (ROS) and inhibit the oxidation of polyunsaturated fatty acids (PUFAs). These droplets protect glia and also neuroblasts from peroxidation chain reactions that can damage many types of macromolecules. The underlying antioxidant mechanism involves diverting PUFAs, including diet-derived linoleic acid, away from membranes to the core of lipid droplets, where they are less vulnerable to peroxidation. This study reveals an antioxidant role for lipid droplets that could be relevant in many different biological contexts.

INTRODUCTION

Stem and progenitor cells drive growth during development, tissue regeneration, and tumorigenesis (Blanpain and Fuchs, 2014; Hanahan and Weinberg, 2011; Slack, 2008). They are regulated by growth factors and other signals from a local microenvironment called the niche (Scadden, 2006). In many different physiological contexts, tissues containing stem cells are exposed to low levels of dietary nutrients or oxygen (Barker, 1995; Dunwoodie, 2009; Pugh and Ratcliffe, 2003). Multiple pathways sense these environmental stresses and trigger responses that can significantly impact upon metabolism (Tower, 2012). Stem, progenitor, and tumor cells tend to utilize proportionately more glycolysis and less oxidative phosphorylation than differentiated cells (Burgess et al., 2014). The low oxygen consumption of glycolytic metabolism appears well matched to the physiological hypoxia of the niche in which many different stem cells reside

(Burgess et al., 2014; Mohyeldin et al., 2010). This hypoxic micro-environment can itself play a key role in regulating the balance between stem cell quiescence, self-renewal, and differentiation. Hypoxic stem cells and their niche are often associated with hypoxia inducible factor (HIF) activity and also with an increase in reactive oxygen species (ROS), both of which can act as signals promoting glycolysis and metabolic reprogramming (Lee and Simon, 2012; Ushio-Fukai and Rehman, 2014). If hypoxia or other forms of oxidative stress induce ROS levels that are high enough to exceed cellular defense mechanisms, then they promote harmful oxidation and peroxidation chain reactions that can damage lipids, proteins, and nucleic acids (Negre-Salvayre et al., 2008). Many types of stem cells therefore synthesize high levels of antioxidants such as glutathione (GSH) and also antioxidant enzymes such as superoxide dismutase (SOD) and catalase (Cat) in order to defend themselves against ROS (Wang et al., 2013).

Neural stem cells are critical for growth of the mammalian CNS during development and also for neuronal turnover in the sub-ventricular zone and dentate gyrus during adulthood (Okano and Temple, 2009; Taverna et al., 2014). In common with other stem cells, they are known to reside in a niche that is hypoxic even when the external environment is normoxic and nutrient rich (Cunningham et al., 2012). It is also well described that neonatal brain size is highly protected or spared from the increased hypoxia and malnutrition that are experienced during intrauterine growth restriction (Barker, 1995; Gruenewald, 1963). How then do these environmental stresses alter the properties of neural stem cells and/or their niche and which, if any, of these changes are adaptive for brain sparing?

The developing *Drosophila* CNS is a useful model for investigating the effects of environmental stresses upon neural stem and progenitor cells. Embryonic and larval neuroepithelia give rise to neural stem cells called neuroblasts, which divide asymmetrically to generate multiple types of neurons and glia (Homem and Knoblich, 2012; Pearson and Doe, 2004; Skeath and Thor, 2003). As with other growing and dividing *Drosophila* cells, larval neuroblasts rely heavily upon glycolytic metabolism (Grant et al., 2010; Homem et al., 2014; Tennessen et al., 2011). Properties of neuroblasts such as their type, number of divisions, and lineage composition differ from region to region within the CNS (Li et al.,

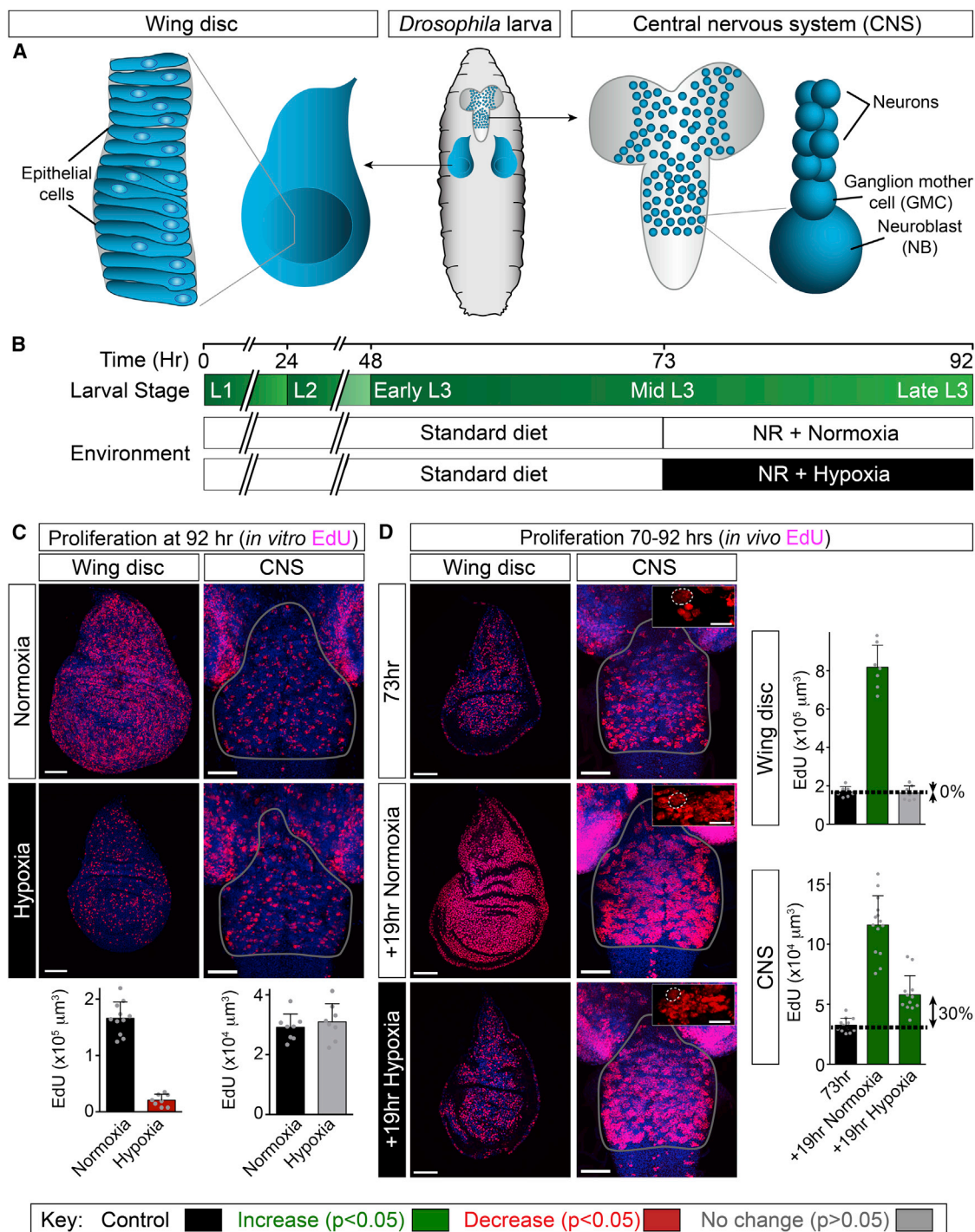


Figure 1. Neuroblast Proliferation Is Spared during Environmental Hypoxia

(A) Cartoon of *Drosophila* larva, showing the developing wing discs and CNS. Most proliferating cells correspond to neuroblasts (NBs) and ganglion mother cells (GMCs) within the CNS and to epithelial progenitors within the wing disc.

(B) Larval development timeline (hours after larval hatching), depicting the three larval instars and the hypoxia (2.5% oxygen) regimen used.

(C) *In vitro* EdU assay shows that cell proliferation in neuroblast lineages (CNS) is more resistant to hypoxia than that of epithelial progenitors (wing disc). After 22 hr of normoxic or hypoxic NR, tissues were dissected and incubated *in vitro* with EdU for 1 hr.

(D) *In vivo* EdU assay indicates that cell proliferation in neuroblast lineages (CNS) is more resistant to hypoxia than that of epithelial progenitors (wing disc). The CNS and wing disc are shown at 73 hr (after 3 hr of normoxic NR with EdU) and also at 92 hr (after a further 19 hr of normoxic or hypoxic NR in the presence of

(legend continued on next page)

2013; Sousa-Nunes et al., 2010). In the central brain and ventral ganglion, neuroblasts undergo two periods of neurogenesis separated by a period of cell-cycle arrest called quiescence. Exit from quiescence (reactivation) occurs during early larval stages and requires dietary amino acids (Britton and Edgar, 1998; Truman and Bate, 1988). Amino acids are sensed by the target of rapamycin (TOR) pathway in the fat body and activate a systemic relay signal that triggers insulin-like peptide (Ilp) expression in glia, an important niche for larval neuroblasts (Chell and Brand, 2010; Colombani et al., 2003; Dumstrei et al., 2003; Sousa-Nunes et al., 2011; Spéder and Brand, 2014). Glial Ilps then activate the insulin-like receptor (InR) in neuroblasts leading to TOR and phosphatidylinositol 3-kinase (PI3K) signaling and re-entry into the cell cycle (Chell and Brand, 2010; Sousa-Nunes et al., 2011). Later during development, the growth of neuroblast lineages becomes largely independent of all dietary nutrients, providing a model for brain sparing (Cheng et al., 2011). At this stage, niche glia express a secreted growth factor, jelly belly (Jeb), regardless of whether dietary nutrients are present or absent. Jeb then activates its receptor anaplastic lymphoma kinase (Alk) in neuroblasts, thus promoting constitutive rather than nutrient-dependent PI3K signaling and growth (Cheng et al., 2011). Similar brain sparing specific to older larvae has also been identified in the *Drosophila* visual system, where the late asymmetric but not the early symmetric divisions of neural progenitors tend to be spared from nutrient restriction (Lanet et al., 2013). Importantly, however, it has not been clear in any region of the *Drosophila* CNS whether neuroblasts are also protected against other environmental stresses such as hypoxia. We now investigate this issue using imaging mass spectrometry, CNS lipidomics, and genetic manipulations specific for the glial niche or the neuroblast lineage. This combined approach identifies a molecular mechanism by which lipid metabolism in a stem cell niche can protect the stem cells themselves from oxidative damage.

RESULTS

The Proliferation of Neuroblast Lineages Is Spared during Hypoxia

Drosophila larvae develop into undersized yet viable adults when starved during the third instar, (Bakker, 1959; Beadle et al., 1938). We previously used a nutrient restriction (NR) protocol in which third instar larvae are transferred from a yeast/glucose/cornmeal to an agarose medium (Cheng et al., 2011). During NR, there is no increase in overall body mass but growth and cell division are spared in CNS neuroblast lineages and also, to a lesser extent, in epithelial progenitors of the wing disc (Cheng et al., 2011; Figure 1A).

We first quantified the effects of hypoxia upon cell proliferation using the thymidine analog 5-ethynyl-2'-deoxyuridine (EdU). This

method measures the proliferation of neuroblast lineages, with glia representing less than 6% of total EdU⁺ cells (Sousa-Nunes et al., 2011; data not shown). Larvae exposed to hypoxia are known to crawl away from yeast paste (Wingrove and O'Farrell, 1999), and we found that, in a 2.5% oxygen environment, they do not ingest significant quantities of food (Figure S1). This behavioral response may normally serve to prevent anoxia by limiting the duration of burrowing bouts in semi-liquid medium, and it can be synchronized by applying artificial hypoxic-normoxic cycles (Movie S1). To assess specifically the effects of low oxygen, rather than reduced feeding, we raised larvae on yeast/glucose/cornmeal until mid-third instar (~70 hr after larval hatching, ALH) and then transferred them to NR medium either in normoxia or hypoxia (Figure 1B). EdU incorporation by thoracic neuroblast lineages and by wing disc progenitors was measured at the end (1 hr in vitro) or throughout (22 hr in vivo) the hypoxic period (Figures 1C and 1D). For both tissues, the percentage of cell proliferation spared in hypoxia (relative to normoxia) is smaller with the in vivo cumulative than with the in vitro endpoint assay. This difference may reflect brief exposure of hypoxic larvae to normoxia prior to dissection and in vitro EdU incubation. Importantly, however, both methodologies clearly demonstrate that cell divisions of neuroblasts are more resistant to hypoxia than those of progenitors in the developing wing.

Hypoxia Induces Lipid Droplets in Subperineurial and Cortex Glia

We next investigated the metabolic specializations underlying the striking hypoxia resistance of neuroblast lineages. Highly active glycolysis in dividing neuroblasts is unlikely to account per se for their selective hypoxia resistance as it is also present in wing discs (de la Cova et al., 2014; Homem et al., 2014; data not shown). However, using neutral lipid stains (oil red O and LipidTOX), we were intrigued to observe lipid droplets in the developing CNS (Figures 2A and 2B). Lipid droplets are cytoplasmic organelles comprising a core of neutral lipids, such as triacylglycerols (TAGs) and cholesteryl esters, surrounded by a phospholipid monolayer associated with lipases and other proteins regulating droplet biogenesis, lipid storage, and release (Kühnlein, 2012; Walther and Farese, 2012; Young and Zechner, 2013). In *Drosophila*, lipid droplets in adipose tissue (known as fat body) express the regulatory protein Lsd-2 (perilipin-2) on their surface (Bickel et al., 2009; Kühnlein, 2011), and we find that this is also the case for droplets in the CNS (Figure 2C). CNS lipid droplets, however, tend to be 1–2 μ m diameter, smaller than many of those in the fat body. On standard food in ambient normoxia, the number of CNS lipid droplets increases during larval development (Figure S2A). This is consistent with a previous lipidomic analysis showing that the late third instar CNS contains significant quantities of TAGs (Carvalho et al., 2012). Lipid droplets in the CNS increase rather than decrease during NR

EdU). CNS insets in (D) show high-magnification views of EdU⁺ cells, large neuroblasts (dotted circle) are already EdU⁺ at 73 hr and contribute to the final EdU volumes.

Histograms in (C) and (D) depict the average volume of EdU⁺ cells incorporated per wing disc or per thoracic CNS. Scale bars, 50 μ m, inset scale bars, 10 μ m. In this and subsequent figures, histograms show the mean, scatterplots show individual data points, and error bars are 1 SD. The key shows that histogram bars are colored according to t tests ($p < 0.05$): significant decrease (red), significant increase (green), or no significant change (gray) from the control (black). See also Figure S1.

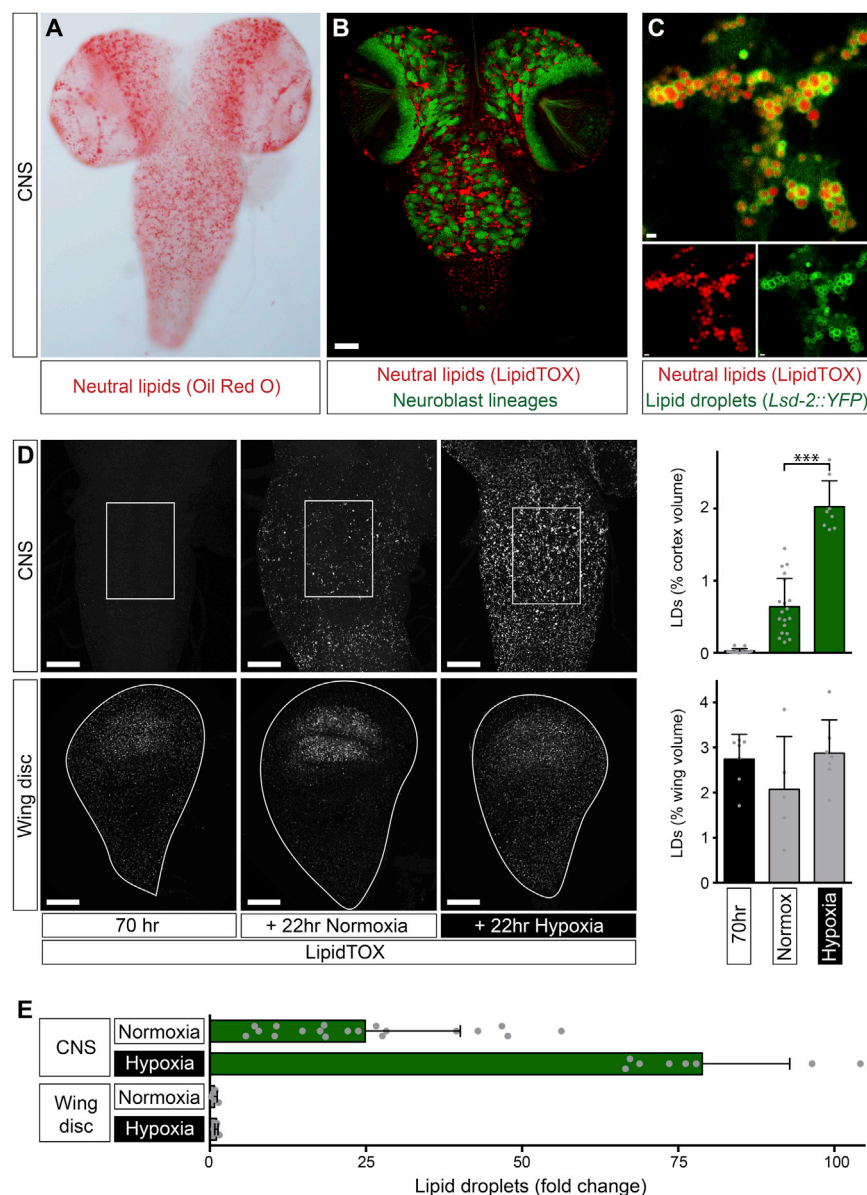


Figure 2. Hypoxia Induces Lipid Droplets in the CNS

(A) Oil red O-stained neutral lipids in the CNS of a late-third instar larva (control genotype: *w¹¹¹⁸*) raised on a standard diet.

(B) LipidTOX stained neutral lipids in the CNS of a late-third instar larva raised on a standard diet. Neutral lipids (red) are prominent in the central brain and ventral nerve cord but do not overlap with neuroblast lineages (green: *nab-Gal4 > CD8::GFP*). Scale bar, 50 μ m.

(C) Neutral lipids (LipidTOX, red) in the thoracic ganglion of the late third instar CNS accumulate in clusters of lipid droplets labeled with *Lsd-2::YFP* (green). Scale bars, 1 μ m.

(D) Lipid droplets (LipidTOX) in the CNS of a mid-third instar (70 hr) larva fed on a standard diet increase during NR normoxia and become more numerous during NR hypoxia. The lipid droplet (LD) content of the wing disc does not change significantly during hypoxia. It is expressed as percentage of volume within the thoracic region of interest of the ventral nerve cord (white boxes, upper row) or the entire wing disc (white outlines, lower row). Scale bars, 50 μ m.

(E) Lipid droplets increase ~25-fold in normoxia and ~80-fold in hypoxia in the CNS but do not significantly change in the wing disc. Note that fold change values in this figure are normalized to the 70 hr start points of Figure 2D.

See also Figure S2.

A panel of GFP and *Lsd-2::GFP* reporters revealed that lipid droplets are primarily found in glial cells but not in neuroblasts or in their neuronal progeny (Figures 3A and 3B). *Lsd-2* localizes to the phospholipid monolayer of the lipid droplet surface where it promotes neutral lipid storage by inhibiting TAG lipolysis (Kühnlein, 2011). Specific knockdown in glia (*repo > Lsd-2 RNAi*) prevents the accumulation of CNS lipid droplets during normal fed development, thus confirming their glial localization (Figures 3C and 3D). Reporters specific for glial subtypes showed that the majority of lipid droplet-containing cells correspond to the sub-perineurial glia of the blood-brain barrier and the cortex glia that enwrap neuroblast lineages (Figures 3A and 3E; Kis et al., 2015). We conclude that CNS lipid droplets mostly localize to the niche (glia) but not to the neural stem cells (neuroblasts).

(with normoxia) and so, in contrast to those in fat body, do not appear to correspond to a nutrient store that is depleted during starvation (Figure 2D; Gutierrez et al., 2007). However, lipid droplets increase during development with a concomitant increase in the distance separating neuroblasts from the oxygen-supplying cells of the tracheal network (Figures S2B–S2D). We then tested directly whether low oxygen tension can induce CNS lipid droplets. Strikingly, exposing larvae to hypoxia induces a 3-fold gain in lipid droplets, over and above the normoxic developmental increase (Figures 2D and 2E). Importantly, this strong induction is CNS specific, as lipid droplets in the wing disc (Fauny et al., 2005; Parra-Peralbo and Culi, 2011) do not increase significantly during hypoxia (Figures 2D and 2E). Hence, lipid droplets accumulate in the CNS during normal development and can also be strongly induced by hypoxia.

Glial Lipid Droplets Sustain Neuroblast Proliferation during Hypoxia

We next determined the metabolic origin of glial lipid droplets using multi-isotope imaging mass spectrometry (MIMS, Steinhäuser et al., 2012). Larvae were raised until mid-third instar on diets containing ^{13}C -labeled glucose or acetate, major carbon sources for lipogenesis (Figure 4A). Glial lipid droplets were then induced with 22 hr hypoxia. MIMS analysis revealed strong

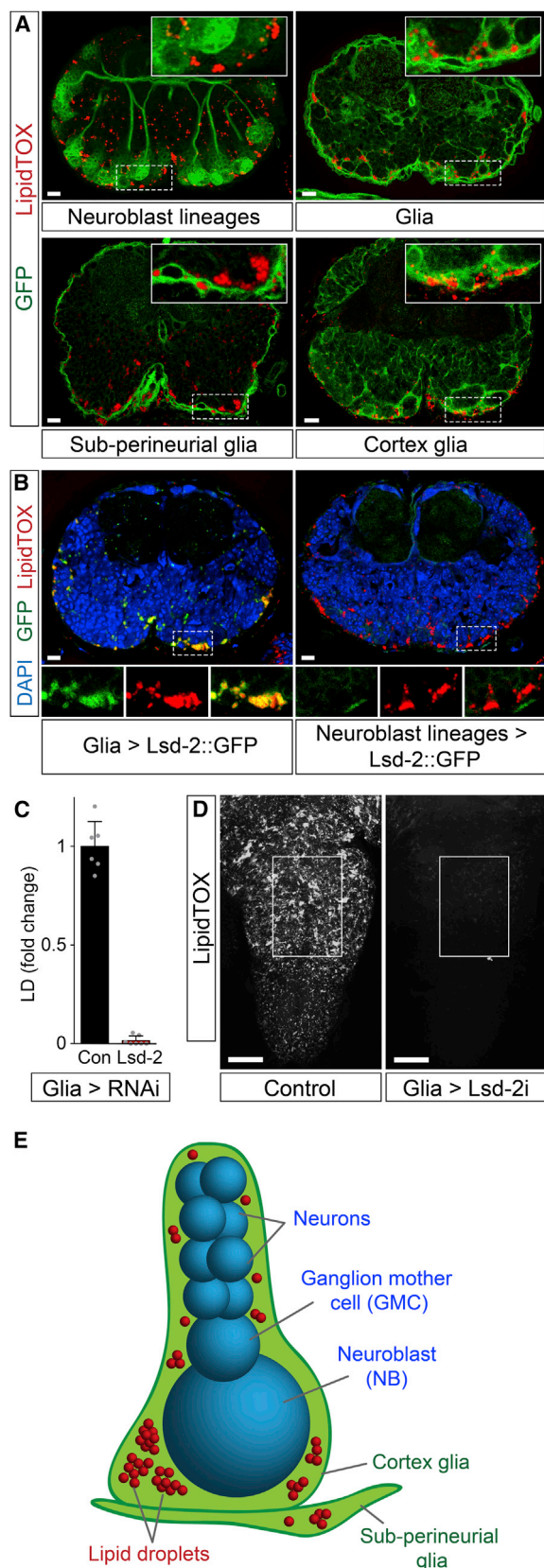


Figure 3. Lipid Droplets Accumulate in Cortex and Subperineurial Glia

(A) Transverse sections (dorsal up) of the thoracic CNS from late-third instar larvae raised on standard diet. Lipid droplets (LipidTOX) are observed nearby, but not within neuroblast lineages (labeled with *nab > CD8::GFP*). They localize to glial cells (*repo > CD8::GFP*) including the cortex glia (*nrl^{GMR37H03} > CD8::GFP*) and the subperineurial glia (*moody > CD8::GFP*). Scale bars, 10 μ m.

(B) Lipid droplets (LipidTOX) localize to glia (*repo > Lsd-2::GFP*) but not to neuroblast lineages (*elav^{C155} > Lsd-2::GFP*). Strong Lsd2::GFP (GFP) signal localizes to the surface of glial lipid droplets but low-intensity signal in glia and neuroblast lineages may indicate localization to other organelles such as the ER (Fauny et al., 2005). Scale bars, 10 μ m.

(C and D) Glial-specific knockdown of Lsd-2 (*repo > Lsd-2 RNAi*) inhibits lipid droplet accumulation in late-third instar larvae raised on standard diet (control: *repo > w¹¹¹⁸*). Scale bars, 50 μ m.

(E) Diagram showing the distribution of lipid droplets (red) within cortex and subperineurial glia (green) surrounding a neuroblast lineage (blue), comprising a single neuroblast, a ganglion mother cell, and multiple neuronal progeny.

^{13}C incorporation from labeled dietary glucose or acetate into the core of hypoxic lipid droplets (Figures 4B and S3A–S3C). Maximal $^{13}\text{C}/^{12}\text{C}$ ratios in tissue cross sections were observed with lipid droplets of 1–2 μ m diameter, consistent with the previous size estimates from confocal microscopy (Figures S3D–S3G). The MIMS analysis demonstrates that de novo fatty acid synthesis contributes neutral lipid cargo to glial lipid droplets. To distinguish whether the relevant de novo lipogenesis is within the glia themselves or in another tissue, glial-specific RNAi was used to knock down six enzymes of TAG biosynthesis (Figure 4A). Lipid droplet induction after 22 hr hypoxia was roughly halved with the knockdown of acetyl-CoA carboxylase (ACC) or a predicted fatty acid synthase (CG3524) (Figure 4C). Fatty acid synthesis is therefore required in a cell-autonomous manner for maximal induction of glial lipid droplets. Given that induction was not completely blocked by knocking down fatty acid synthetic enzymes, we also tested whether there is an additional contribution from dietary fatty acids. Linoleic acid (C18:2) is an omega-6 polyunsaturated fatty acid (PUFA) that is derived from the diet and cannot be synthesized by *Drosophila* (Figure 4A). It is one of the major fatty acids in our standard *Drosophila* diet and remains largely intact following ingestion. Thus, larvae raised on a diet containing ^{13}C -linoleic acid retain ~80% of the ^{13}C label incorporated into CNS fatty acids as linoleate (Figure S3H). MIMS analysis of these larvae revealed strong ^{13}C incorporation in hypoxia-induced CNS lipid droplets (Figure 4B). Together with the previous results, this indicates that dietary uptake of fatty acids (such as linoleate) and fatty acid synthesis both contribute to glial lipid droplets.

Partial decreases in hypoxia-induced lipid droplets were observed for glial knockdowns of the *Drosophila* orthologs of genes encoding enzymes converting fatty acids into lysophosphatidic acids (GPAT, CG3209, one of three glycerol-3-phosphate acyltransferases) or lysophosphatidic acids into phosphatidic acids (AGPAT, CG4753, one of four 1-acyl-sn-glycerol-3-phosphate acyltransferases) (Figure 4A; for enzyme predictions, see Wilfling et al., 2013). More strikingly, hypoxia-induced lipid droplets were reduced 95%–100% with knockdown of the enzymes converting phosphatidic acids into diacylglycerols (Lipin, 3-sn-phosphatidate phosphohydrolase),

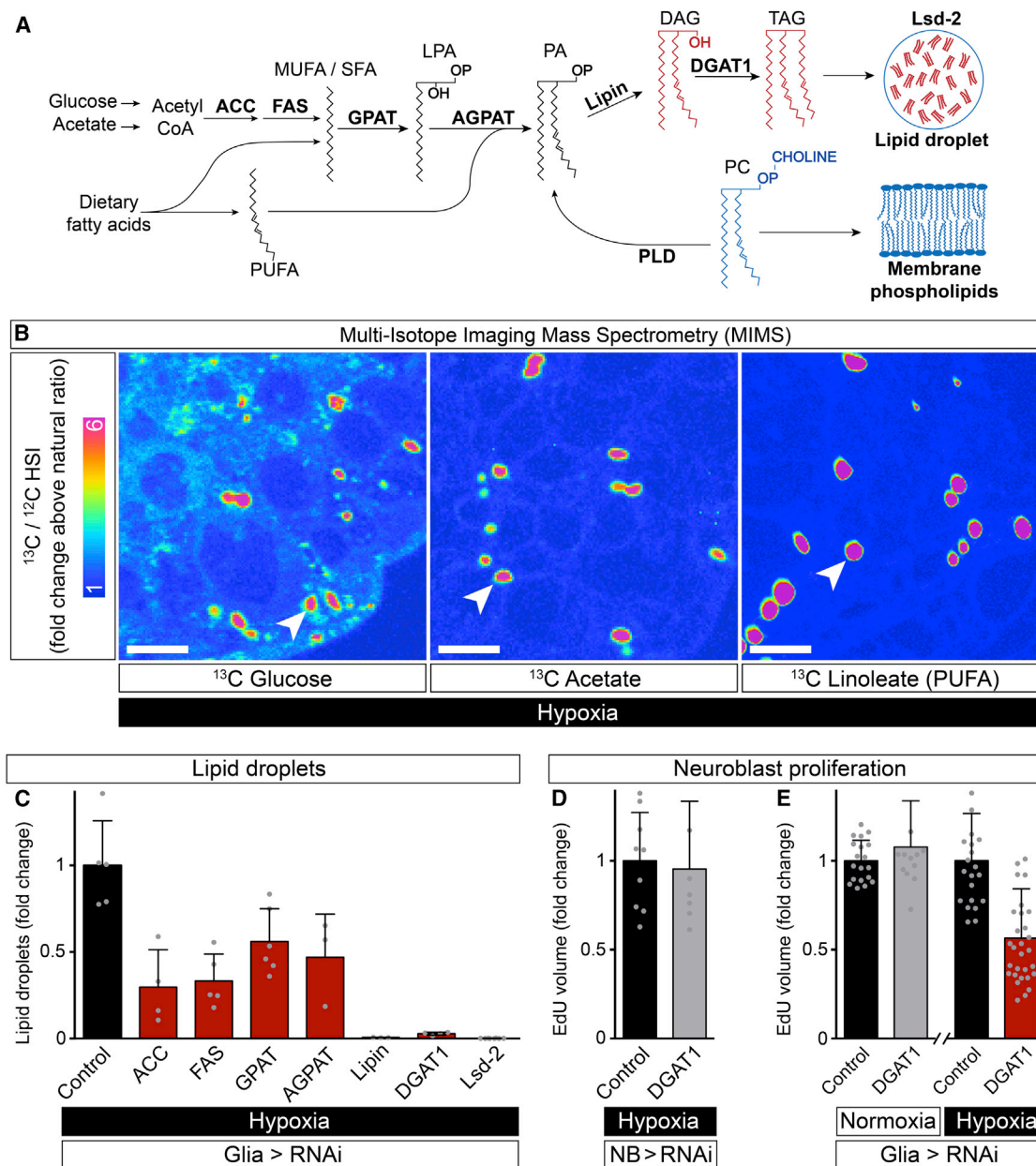


Figure 4. Neuroblast Proliferation during Hypoxia Requires Glial TAG Synthesis

(A) TAG synthetic pathway showing the enzymes (bold, see text for details) knocked down by cell-type-specific RNAi in this study. Dietary fatty acids that are saturated (SFA), monounsaturated (MUFA), or polyunsaturated (PUFA) can be incorporated into the pathway at multiple steps, but, for clarity, single entry points are depicted.

(B) Multi-isotope imaging mass spectrometry (MIMS) of glial lipid droplets induced by hypoxia in larvae raised on diets containing $1\text{-}^{13}\text{C}$ -glucose, $1\text{-}^{13}\text{C}$ acetate, or ^{13}C -U-linoleate (PUFA). Hue saturation intensity (HSI) images show that high $^{13}\text{C}/^{12}\text{C}$ ratios are selectively detected in CNS lipid droplets (arrowhead shows one example). Scale bars, 5 μm .

(C) Quantification of hypoxia-induced lipid droplets (fold change at endpoint versus *repo* > w^{1118} control) with glial-specific gene knockdowns (*repo*>RNAi) of TAG synthetic genes.

(D and E) Quantification of neuroblast proliferation (in vivo EdU assays), during normoxia or hypoxia with neuroblast lineage (*nab-Gal4*) or glial (*repo-Gal4*)-specific DGAT1 knockdown (> *DGAT1* RNAi, control: > w^{1118}). Pooled data are represented as mean \pm SD.

See also Figure S3.

diacylglycerols into TAGs (DGAT1, a diacylglycerol O-acyltransferase), or with knockdown of Lsd-2, which promotes TAG storage in lipid droplets (Figures 4A and 4C). The RNAi results

together suggest that TAGs are the predominant neutral lipids stored in glial droplets. They also demonstrate clearly that enzymes converting fatty acids into TAGs are required in glia for

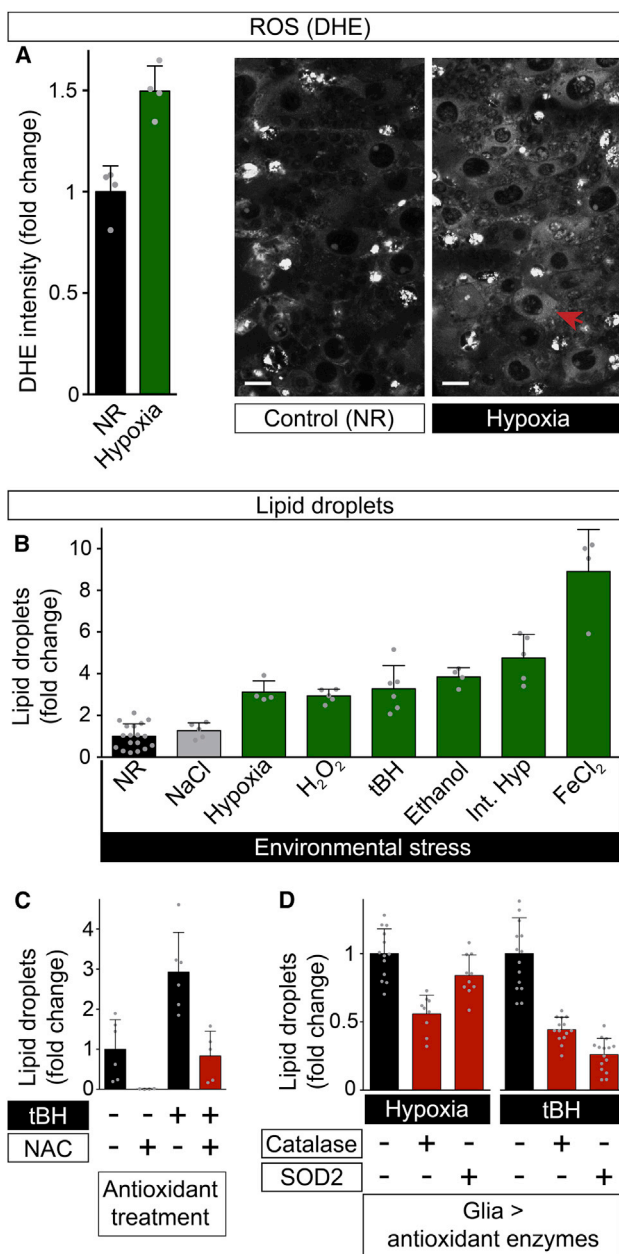


Figure 5. ROS Induce CNS Lipid Droplets

(A) Hypoxia significantly increases ROS (measured by DHE oxidation) in many cells of the larval CNS including neuroblasts (red arrow marks one of the neuroblasts). Scale bars, 10 μ m.

(B) CNS lipid droplets increase (fold change relative to NR endpoint) in response to 22 hr of in vivo exposure to the following pro-oxidants: chronic hypoxia (2.5% oxygen), H₂O₂ (0.5% v/v), tBH (100 mM), ethanol (10% v/v), intermittent hypoxia (44 cycles of 10 min normoxia: 10 min anoxia) or FeCl₂ (100 mM). Note that 22 hr of osmotic stress (1.25 M NaCl) does not significantly increase lipid droplets.

(C) Dietary supplementation with the antioxidant N-acetyl cysteine (100 mM) inhibits lipid droplet induction by NR or tBH (100 mM).

(D) Glial-specific overexpression of Catalase (*repo>Cat*) or Superoxide Dismutase 2 (*repo>SOD2*) inhibits lipid droplet induction by hypoxia (2.5% oxygen) or tBH (100 mM) (control genotype: *repo > w¹¹¹⁸*).

See also Figure S4.

the hypoxic induction of lipid droplets. Importantly, there is a strong requirement for DGAT1, one of two enzymes catalyzing the final and only dedicated reaction in TAG synthesis, which is required for droplet biogenesis from the endoplasmic reticulum (Wilfling et al., 2014).

To test whether the hypoxic induction of glial lipid droplets is functionally linked to neuroblast proliferation, we overexpressed YFP::Pros to drive dividing neuroblasts into premature differentiation (Maurange et al., 2008). This failed to inhibit glial lipid droplets, indicating that neuroblast proliferation is not required for glial lipid droplet induction during hypoxia (Figures S4A and S4B). The preceding RNAi results also allowed us to test for the reciprocal regulatory relationship, namely, that lipid droplets are required for neuroblast proliferation. Control experiments indicated no significant effects upon neuroblast proliferation if DGAT1 is knocked down in neuroblast lineages during normoxia or hypoxia, and also if DGAT1 is knocked down in glia during normoxia (Figures 4D and 4E; data not shown). However, glial-specific DGAT1 knockdown during hypoxia approximately halved neuroblast proliferation (Figure 4E). Together with the previous results, this finding demonstrates that the synthesis and storage of TAGs, as lipid droplets in niche glia, are required to sustain neuroblast divisions during hypoxia but not during normoxia.

Reactive Oxygen Species Induce Glial Lipid Droplets

To investigate the mechanism by which hypoxia induces glial lipid droplets, we first tested the HIF pathway. However, no substantial stabilization of a HIF::GFP fusion protein was observed in glia during hypoxia, and *sima* homozygotes lacking *Drosophila* HIF-1 activity retained the ability to accumulate lipid droplets during hypoxia (Figures S4C and S4D). We next considered alternative induction mechanisms. Hypoxia can markedly increase the levels of reactive oxygen species (ROS) generated by the mitochondrial electron transport chain, which, in turn, can lead to oxidative stress and cellular damage (Bleier and Dröse, 2013; Chandel et al., 1998; Olsen et al., 2013). Even in normoxic controls, the superoxide reactive dye dihydroethidium (DHE) detected more ROS in neuroblasts than in many other cells of the CNS (Figure S4E). Following hypoxia, ROS were moderately increased throughout the CNS, with higher levels detectable in neuroblasts (Figure 5A). We therefore tested whether intermittent hypoxia and a variety of chemical pro-oxidants known to increase ROS (hydrogen peroxide, tert-butyl hydroperoxide (tBH), ethanol, or iron (II) chloride) would be able to induce lipid droplets. All of these pro-oxidants induced glial lipid droplets at least as strongly as chronic hypoxia, whereas osmotic stress (1.25 M NaCl) did not (Figure 5B). The antioxidant N-acetyl cysteine (NAC) is a ROS scavenger and it replenishes the reduced glutathione pool (Atkuri et al., 2007). When added to the substrate, NAC inhibited both the moderate and strong increases of glial lipid droplets observed during NR and tBH stress, respectively (Figure 5C). In addition, glial-specific overexpression of the antioxidant enzymes catalase (Cat) or superoxide dismutase 2 (Sod2) inhibited hypoxic or tBH induction of glial lipid droplets (Figure 5D). Together, the chemical and genetic manipulations demonstrate that ROS are an important stimulus for inducing glial lipid droplets during normal development and in the presence of environmental pro-oxidants.

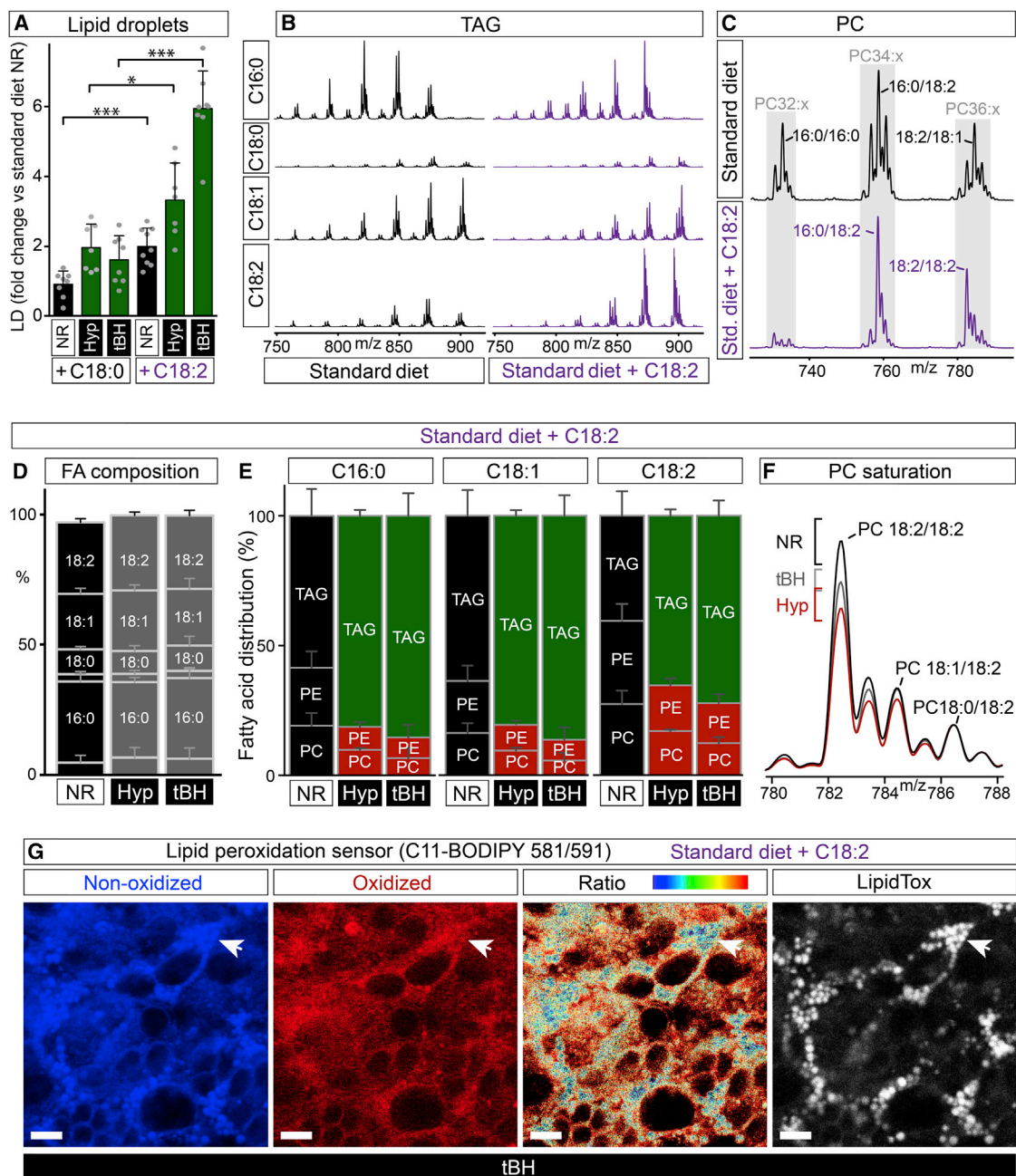


Figure 6. PUFAs Redistribute from Phospholipids to TAGs during Oxidative Stress

(A) Lipid droplets following NR, hypoxia (2.5% oxygen), or tBH (100 mM) are more abundant on a diet supplemented with 20 mM linoleate (C18:2) than with 20 mM stearate (C18:0).

(B) Neutral loss scans indicate that palmitate (C16:0), stearate (C18:0), oleate (C18:1), and linoleate (C18:2) are the major fatty acids present in CNS TAGs following NR. Dietary 20 mM linoleate (C18:2) leads to a relative enrichment of this PUFA in the TAG pool. X values are m/z, and Y values are relative signal intensities, scaled to the largest C16:0 peak.

(C) Positive ion scans indicate the major PC mass envelopes (gray boxes) in the CNSs of NR larvae. Dietary 20 mM linoleate (C18:2) leads to a relative decrease in PC 16:0/16:0 and a relative enrichment of PC species containing C18:2. X values are m/z and Y values are relative signal intensities, scaled to the PC 16:0/18:2 peak.

(D) Total FA composition (by GC-MS) of larvae raised on 20 mM linoleate diet. Neither hypoxia (n = 3) nor 100 mM tBH (n = 3) lead to major changes, compared to NR (n = 3), in the relative abundance of the major fatty acids (top to bottom: C18:2, C18:1, C18:0, C16:1, C16:0, and C14:0).

(E) Relative distribution of major fatty acids (C16:0, C18:1, and C18:2) in the TAG, PC, and PE pools of CNSs from larvae on 20 mM linoleate diet. Hypoxia (n = 3) and 100 mM tBH (n = 4) both lead to a strong decrease, compared to NR (n = 3), in the proportion of all three major fatty acids in PCs and PEs, with a correspondingly strong increase in their proportions in TAG.

(legend continued on next page)

PUFAs Redistribute from Phospholipids to TAGs during Oxidative Stress

ROS attack the carbon-carbon double bonds in linoleic acid and other PUFAs, inducing lipid peroxidation chain reactions that can damage membrane phospholipids as well as other macromolecules (Negre-Salvayre et al., 2008). In the presence of oxidative stress (hypoxia or tBH), glial lipid droplets are induced more strongly when the diet is supplemented with linoleic (C18:2) than with stearic (C18:0) acid (Figure 6A). Mass spectrometry of isolated CNSs indicates that dietary supplementation with linoleic acid increases the proportion of this PUFA in the TAG, phosphatidylcholine (PC), and phosphatidylethanolamine (PE) pools (Figures 6B and 6C; data not shown). On the high-PUFA diet, neither hypoxia nor tBH detectably alter the total fatty acid profile of the CNS (Figure 6D). Both oxidative stresses do, however, produce a large decrease in the proportion of CNS linoleate in membrane phospholipids (PC and PE), with a correspondingly large increase in the proportion in TAG (Figure 6E). This redistribution from phospholipids to TAG is not selective for linoleate as it is also observed with the abundant monounsaturated and saturated fatty acids oleate (C18:1) and palmitate (C16:0) (Figure 6E). Together with our previous confocal microscopy and MIMS analyses, the mass spectrometry strongly suggests that oxidative stress leads to a bulk redistribution of fatty acids (including linoleic acid) from membranes to lipid droplets.

To identify how fatty acids are redistributed from CNS membranes to lipid droplets, we developed a simplified *in vitro* model that minimizes any input from fatty acid uptake or synthesis. Explanted CNSs were cultured in a saline medium lacking all lipids, sugars, and amino acids. This stress is sufficient to induce lipid droplets in an ACC-independent manner, but they are much smaller than those formed *in vivo* (Figures S5A and S5B). Nevertheless, droplet accumulation *in vitro* does require DGAT1 and Lsd-2, as it does *in vivo* (Figure S5B). Moreover, we found that the enzymes catalyzing the conversions of phosphatidylcholine (PC) into phosphatidic acid (PA) and PA into diacylglycerol (phospholipase D (PLD) and Lipin, respectively) are both required in glia for lipid droplet induction *in vitro* and *in vivo* (Figures 4A, S5B and S5C). Together, the *in vitro* and *in vivo* genetic analyses suggest that transfer of fatty acyl chains from membrane phospholipids to TAGs, via a PLD/Lipin/DGAT1 pathway, contributes to the biogenesis of glial lipid droplets during oxidative stress.

Glial Lipid Droplets Protect Dividing Neuroblasts from PUFA Peroxidation

On the high PUFA diet, oxidative stress from tBH or hypoxia selectively depletes linoleic acid (relative to monounsaturated and saturated fatty acids) from the PC pool of the CNS (Figure 6F). On a standard diet in the presence of iron (II), a strong

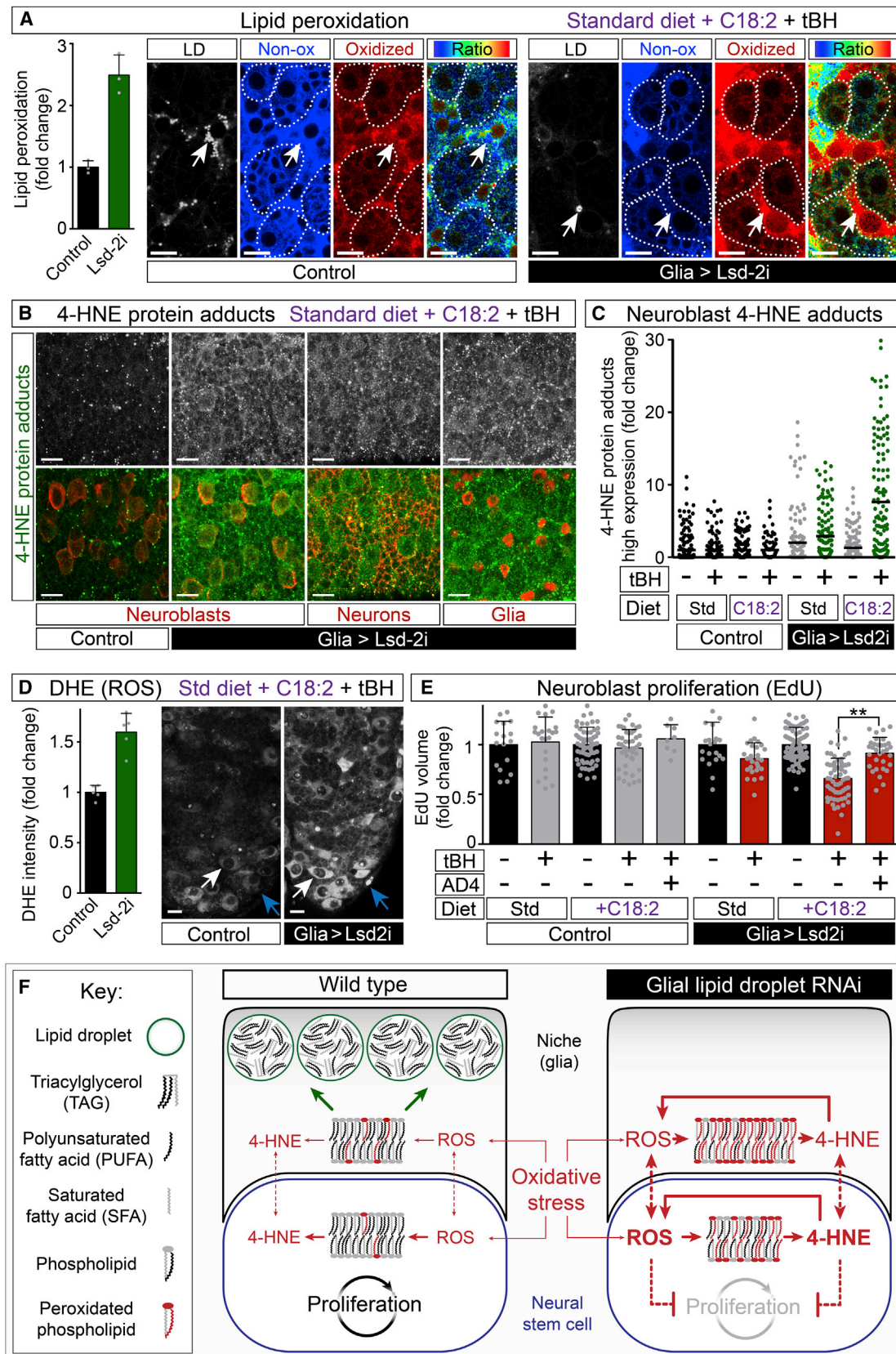
pro-oxidant (Olsen et al., 2013; Figure 5B), selective depletion of linoleate from PCs was also observed (Figure S5D). Although linoleic acid is selectively lost from the PC pool during oxidative stress, we failed to detect its selective increase (or decrease) in the TAG fraction during iron, tBH, or hypoxic stress (Figures S5E and S5F). Hence, the loss of linoleate from PCs may not reflect its selective remobilization to lipid droplets but could result from its selective susceptibility (over more saturated fatty acids) to ROS-induced peroxidative damage. Indeed, tBH stress on a high linoleate diet does lead to increased PUFA peroxidation, as detected by a ratiometric fluorescent sensor (C11-BODIPY 581/591, Figure 6G). Strikingly, the ratio of oxidized to non-oxidized sensor in the CNS is much higher in cell membranes than in lipid droplets. Importantly, the PUFA peroxidation ratio of CNS membranes increases even further if lipid droplet accumulation is abrogated by Lsd-2 knockdown (Figure 7A). Even though the knockdown was glial specific and a very strong increase in PUFA peroxidation was observed in glial membranes, a moderate increase in PUFA peroxidation was also detected in the membranes of neuroblast lineages. This indicates that glial lipid droplets exert a non-cell autonomous and protective effect upon neuroblast lineages. Together with the mass spectrometry, the fluorescent sensor provides evidence that PUFAs like linoleic acid are more susceptible to peroxidation when incorporated into membrane phospholipids than when stored as TAGs in the core of lipid droplets.

PUFA peroxidation can also lead to the damage of non-lipid macromolecules. For example, peroxidation chain reactions of omega-6 PUFAs such as linoleic acid generate 4-hydroxy-2-nonenal (4-HNE), an unsaturated aldehyde that forms covalent adducts with proteins (Uchida, 2003). Following *in vitro* incubations with exogenous 4-HNE, such protein adducts can be detected by immunocytochemistry in most/all cells of the CNS, including neuroblasts (Figure S6A). Consistent with studies in other biological contexts (Uchida, 2003), 4-HNE induced ROS throughout the CNS but, as seen previously during hypoxia and even during normal development, levels were higher in cells of a large size characteristic of neuroblasts (Figure S6B). 4-HNE also markedly decreased the proliferation of neuroblasts, but this was efficiently rescued with the antioxidant N-acetylcysteine amide (AD4) (Figures S6C and S6D). These *in vitro* CNS experiments reveal that 4-HNE and ROS induce each other and also that high levels of one or both are deleterious for neuroblast proliferation.

We next determined *in vivo* whether glial lipid droplets function not only to defend membrane lipids against peroxidation, but also to protect proteins from 4-HNE damage. For the control genotype, no obvious increase in 4-HNE protein adducts was observed with tBH stress on standard or high PUFA diets (Figures 7B and 7C). In contrast, for glial Lsd-2 knockdown, tBH

(F) Positive ion scans of PC 18:x/18:x from the CNSs of larvae on a 20 mM linoleate diet. The ratio of 18:2/18:2 to 18:0/18:2 is decreased by hypoxia ($n = 3$, mean = 4.03, SD = 0.43, $p = 0.008$) and tBH ($n = 3$, mean = 4.84, SD = 0.3, $p = 0.09$) relative to the NR control ($n = 3$, mean = 5.7, SD = 0.62). For each treatment group, the mean spectrum is normalized to the 18:0/18:2 peak (m/z 786) of NR, and the maximum/minimum peak heights at m/z 782 (brackets) indicate variation between biological replicates.

(G) CNS lipid peroxidation following 100 mM tBH treatment of larvae raised on a 20 mM linoleate diet. Co-incubation with the PUFA peroxidation sensor (C11-BODIPY 581/591) and LipidTOX 633 shows that the ratio of oxidized: non-oxidized sensor is lower in CNS lipid droplets (white arrow) than in cell membranes. Scale bars, 10 μ m. See also Figure S5.



(legend on next page)

stress increased 4-HNE protein adducts significantly in neuroblasts, and this increase was further augmented by a high PUFA diet (Figures 7B and 7C). Cell-type markers indicated that Lsd-2 knockdown increases 4-HNE protein adducts not only in neuroblasts, but also in glia and in neurons (Figure 7B). Under these *in vivo* conditions, glial lipid droplet knockdown led to increased ROS throughout the CNS, with maximal levels in neuroblasts (Figure 7D). These findings demonstrate that glial lipid droplets are required to defend the proteins of neuroblasts and other neural cells against 4-HNE damage from even a moderate oxidative challenge.

Finally, we tested whether a high PUFA diet influences the ability of glial lipid droplets to protect neuroblast proliferation. With intermittent hypoxia on a high PUFA diet, a slight decrease in neuroblast proliferation was observed, and this was exacerbated with glial Lsd-2 knockdown (Figure S6E). The milder oxidative stress of tBH on a high PUFA diet (or on a standard diet) did not significantly alter proliferation nor detectably increase CNS apoptosis in larvae of the control genotype (Figure 7E; data not shown). In larvae with glial knockdown of Lsd-2, however, proliferation was decreased slightly by tBH and much more strongly by the combination of tBH with high dietary PUFA (Figure 7E). This deficit in proliferation could be significantly rescued by dietary supplementation with the antioxidant AD4. Hence, when the diet is rich in PUFAs, glial lipid droplets play a particularly critical role in sustaining the proliferation of neuroblasts during oxidative stress.

DISCUSSION

This study reveals that the proliferation of *Drosophila* neural stem cells is defended against oxidative stress and identifies lipid droplets in niche glia as a critical element of the protective mechanism. During oxidative stress, PUFAs and other fatty acids are redistributed from membrane phospholipids to lipid droplet TAGs. Unlike cell membranes, the lipid droplet core provides a

protective environment that minimizes PUFA peroxidation chain reactions and limits ROS levels. This helps to safeguard not only the glial cells of the niche but also the neighboring neural stem cells and their neuronal progeny. We now compare the protective roles of the niche glia during oxidative stress and starvation, discuss the lipid droplet antioxidant mechanism, and speculate upon its wider implications.

Niche Glia Protect Neuroblasts from Oxidative Stress and Nutrient Deprivation

For both NR and oxidative stress, the protection of neuroblast divisions involves niche glia. We previously showed that neural growth and proliferation are spared during NR by constitutive glial secretion of Jeb, a ligand for Alk/PI3-kinase signaling in neuroblasts (Cheng et al., 2011). In the present study, we find that neuroblast proliferation is also spared during oxidative stress and that this requires the glial activity of proteins involved in TAG metabolism such as DGAT1 and Lsd-2. These enzymes promote the accumulation of PUFA and other fatty acids as lipid droplet TAGs in cortex and subperineurial glia. TAG synthesis is required in these niche glia for sustaining neuroblast proliferation during oxidative stress but this is not the case during NR or fed conditions. Hence, although niche glia are critical for protecting neighboring neuroblasts from nutrient deprivation and from oxidative stress, in each case a different molecular mechanism is involved. Presumably both glial mechanisms work together to give neuroblasts the remarkable ability to continue dividing even when oxidative stress is combined simultaneously with NR.

A High PUFA Diet Increases the Vulnerability of Neuroblasts to ROS Damage

We found that a high PUFA diet makes neural stem cells and their glial niche more vulnerable to ROS-induced damage. Omega-3 and omega-6 PUFAs are obtained from the diet as both *Drosophila* and humans lack the key desaturases required to

Figure 7. Lipid Droplets Protect Dividing Neuroblasts from PUFA Peroxidation

(A) Lipid peroxidation following low-dose (10 mM) tBH treatment for 20 hr of larvae raised on a 20 mM linoleate (C18:2) diet. Co-incubation with the PUFA peroxidation sensor (C11-BODIPY 581/591) and LipidTOX 633 shows an increase in the ratio of oxidized: non-oxidized sensor in neuroblast lineages (dotted outlines) and also in glia following the knockdown of lipid droplets (*repo* > *Lsd-2 RNAi*, control genotype: *repo* > *w¹¹¹⁸*). With *repo* > *w¹¹¹⁸*, there is a low peroxidation ratio in LipidTOX⁺ droplets (arrow), but, with *repo* > *Lsd-2 RNAi*, there is a high peroxidation ratio in rare abnormal LipidTOX⁺ puncta (arrow). The same two genotypes are used in (B)–(E). Scale bars in this and subsequent panels, 10 μ m.

(B) Glial lipid droplet knockdown increases 4-HNE protein adducts in many cells of the CNS, including neuroblasts (marked with anti-Miranda), neurons (marked with Anti-Neurotactin), and glia (marked with anti-Repo).

(C) Quantification of 4-HNE protein adducts in neuroblasts. Glial lipid droplet knockdown in the presence of 10 mM tBH (+ tBH), but not in NR controls (– tBH), leads to a significant increase in 4-HNE protein adducts. This is further increased when larvae are raised on a 20 mM linoleate diet. Mann-Whitney test: gray $p > 0.1$, green $p < 0.001$ relative to the appropriate genotype control.

(D) ROS increase following glial lipid droplet knockdown. Oxidized DHE levels increase significantly throughout the CNS (histogram) with elevated levels in glia (blue arrows) and in neuroblasts (white arrows).

(E) Neuroblast proliferation is significantly inhibited by 10 mM tBH if glial lipid droplets are knocked down. Inhibition is stronger when larvae are raised on a 20 mM linoleate diet and can be significantly rescued by dietary supplementation with 40 μ g/ml AD4. In the control genotype, tBH and 20 mM dietary linoleate do not significantly decrease neuroblast proliferation. Pooled data are represented as mean \pm SD.

(F) PUFA protection model for the antioxidant role of lipid droplets. The neural stem cell niche (glia) and the neural stem cell (neuroblast) are depicted in the presence of oxidative stress, in a CNS that is wild-type (left) or RNAi knockdown for a glial lipid droplet gene (right). Oxidative stress stimulates the biogenesis of glial lipid droplets that protect vulnerable PUFAs in the TAG core from ROS-induced peroxidation. Peroxidation of omega-6 PUFAs produces 4-HNE, which, in turn, can generate more ROS, leading to PUFA chain reactions that damage membrane lipids and other macromolecules such as proteins. Following oxidative stress, the proportion of total PUFAs and other fatty acids increases in the TAG pool, relative to the phosphatidylcholine pool. This may correspond to bulk redistribution of lipids from membranes to lipid droplets, which would decrease the amount of membranes and thus minimize the fraction of PUFAs that are susceptible to ROS-induced peroxidation. The mechanism by which ROS and/or 4-HNE inhibit stem cell proliferation and the molecule(s) communicating between the niche and the stem cell are not yet clear (dotted arrows).

See Discussion for details. See also Figure S6.

synthesize them. These PUFAs are essential fatty acids for humans but they do not appear to be required for *Drosophila* survival (Rapport et al., 1983). Larvae do, however, consume significant quantities of the omega-6 PUFA linoleic acid and increasing its concentration in the diet makes neuroblasts and other developing neural cells more susceptible to damage from ROS-induced lipid peroxidation chain reactions. Glial lipid droplets make an important contribution toward minimizing this ROS damage in the CNS by storing linoleic acid and protecting it from peroxidation. Saturated and monounsaturated fatty acids are much less vulnerable than PUFAs to peroxidation, but they too accumulate in glial lipid droplets during oxidative stress. Whether these fatty acids contribute to the protection of PUFAs in lipid droplets is not yet clear but, in the membranes of cancer cells, they are thought to decrease ROS damage by diluting PUFAs in the phospholipid pool (Rysman et al., 2010). It may also be important to avoid excess saturated fatty acids in membrane phospholipids, and the scavenging of extracellular mono-unsaturated lysophospholipids provides one way of preventing this (Kamphorst et al., 2013).

An Antioxidant Role for the Lipid Droplet

It is becoming increasingly clear that lipid droplets mediate cellular functions other than fat storage and mobilization relevant for energy homeostasis. For example, lipid droplets can participate in protein degradation, histone storage, viral replication, and antibacterial defense (Anand et al., 2012; Walther and Farese, 2012). Our study now reveals an additional role for lipid droplets as an antioxidant organelle, defending membrane PUFAs from damage by ROS-induced peroxidation. We also found that lipid droplets play a wider antioxidant role, protecting cellular proteins from reactive peroxidation products such as 4-HNE. The results presented here support a PUFA protection model for the antioxidant mechanism of lipid droplets (Figure 7F). Central to this model are five key findings. First, DGAT1 and Lsd-2 knockdowns and AD4 rescues show that glial lipid droplets act during oxidative stress to minimize ROS, PUFA peroxidation, and 4-HNE damage and also to sustain neuroblast proliferation. Second, lipid peroxidation sensor and tandem mass spectrometry experiments argue that the major diet-derived PUFA, linoleic acid, is less vulnerable to peroxidation in lipid droplet TAGs than in membrane phospholipids. Third, Catalase or SOD2 overexpression or NAC supplementation provides evidence that ROS (directly or indirectly) induce lipid droplets during oxidative stress. Fourth, MIMS, tandem mass spectrometry, PLD knockdown, and CNS explants demonstrate that saturated, monounsaturated, and polyunsaturated fatty acids redistribute from membrane phospholipids to lipid droplet TAGs during oxidative stress. And fifth, in vitro and in vivo experiments together show that pro-oxidants not only stimulate widespread ROS and PUFA peroxidation but also that 4-HNE, a product of PUFA peroxidation, is itself capable of increasing ROS, particularly in neuroblasts. This identifies a positive feedback loop in the CNS between ROS, PUFA peroxidation, and 4-HNE but also suggests that it will be difficult to pinpoint the molecule(s) propagating oxidative damage between glia and neuroblasts.

Our findings raise the important issue of why PUFAs are more efficiently protected from peroxidation in the core of lipid droplets

than in membranes. Future biophysical investigations will be needed to resolve this but it is likely that lipid droplets provide a means to segregate reactive PUFAs away from aqueous pro-oxidants and lipid peroxidation chain reactions at membrane-aqueous interfaces. Studies of oil-and-water emulsions indicate that TAGs in small lipid droplets are less mobile and thus better protected from peroxidation chain reactions than those in larger lipid droplets (Nakaya et al., 2005). Interestingly, lipid droplets in glia are smaller, on average, than those found in the major lipid storage depot of the fat body. Given this, and that their biogenesis in glia requires DGAT1, it is interesting that a previous study found that DGAT1 and DGAT2 are involved in the formation of small versus large lipid droplets, respectively (Wilfling et al., 2013).

Lipid droplets have been observed in many different cell types challenged with hypoxia, ischemia, or various metabolic imbalances. For example, mitochondrial protein knockdowns in *Drosophila* photoreceptor neurons generate ROS and induce lipid droplets in neighboring pigment and epithelial glia (Liu et al., 2015). In this mutant context, either neuronal or glial-specific overexpression of lipases can decrease lipid droplets and ameliorate photoreceptor degeneration. In a mammalian example, hypoxia induces lipid droplets in glioblastoma and breast cancer cells via a pathway involving HIF-1 α (Bensaad et al., 2014). This hypoxic induction mechanism in cancer cells appears different from the HIF-1 independent pathway that we have identified in the *Drosophila* neural stem cell niche. It is nevertheless intriguing that, in both contexts, inhibiting lipid droplet formation increases ROS toxicity and impairs cell proliferation. It will therefore be important in future to determine whether lipid droplets also play antioxidant roles in other biological contexts.

EXPERIMENTAL PROCEDURES

Larval Environmental Manipulations and EdU Labeling

Larvae at 0–6 hr after larval hatching (ALH) were raised, 20 per vial, at 25°C on standard cornmeal/yeast/agar food (composition in Gutierrez et al., 2007) until mid-third instar (66–72 hr ALH), then subjected to environmental stresses. For NR, larvae were floated from standard diet using 30% glycerol/PBS and then transferred to 0.5% low melting-point agarose. N-acetyl cysteine (pH 7.2) and/or pro-oxidants were added to agarose at 38°C, before cooling. Stearic or linoleic acid (Sigma-Aldrich) was supplemented to 20 mM in standard food. Hypoxia used an O₂ Cabinet (Coy Lab Products) calibrated to 0% and 20.9% oxygen, with 2.5% oxygen maintained by automated nitrogen injection. Intermittent hypoxia used alternate injections of nitrogen and air over multiple cycles. For the in vitro EdU assay, dissected tissues were incubated at 25°C for 1 hr in PBS with 10 μ M EdU. For the in vivo EdU assay, larvae were transferred to 0.5% low melting-point agarose with 100 μ M EdU for 3 hr and stresses then applied in the presence of EdU for a further 19 hr.

Tissue Imaging and Analysis

Tissues were fixed and stained for EdU with the Click-iT EdU Alexa Fluor 555 Imaging Kit (Molecular Probes) and for neutral lipids using oil red O or LipidTOX (Molecular Probes). Live ex vivo imaging was used to detect oxidized DHE (Invitrogen) and peroxidized lipid (C11-BODIPY 581/591, Invitrogen). See Supplemental Experimental Procedures for details.

Lipid Analysis by MIMS, Tandem MS, and GC-MS

For Multi-Isotope Imaging Mass Spectrometry (MIMS), larvae were fed diets containing 1-¹³C Glucose, 1-¹³C acetate, or ¹³C-U-linoleic acid (Cambridge Isotope Laboratories), and after hypoxia (2.5% oxygen for 22 hr) CNSs were dissected, fixed, embedded in Technovit 7100 (Heraeus Kulzer), and

sectioned at 2 μm . Analysis was performed on 30- to 50- μm regions of interest with simultaneous imaging of multiple ions, including $^{12}\text{C}^{14}\text{N}$, ^{12}C , and ^{13}C . Tandem mass spectrometry (nanospray infusion) and gas chromatography mass spectrometry were performed on lipid extracts of 6–15 CNSs. See Supplemental Experimental Procedures for details.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.020>.

AUTHOR CONTRIBUTIONS

A.P.B. and A.P.G. designed the experiments and wrote the paper. A.P.B. conducted *Drosophila* experiments, MIMS sections were prepared by E.M.A.H. and analyzed by C.G. and C.P.L., and lipid extracts were analyzed by tandem mass spectrometry (G.K. and A.D.P.) and GC-MS (J.I.M.).

ACKNOWLEDGMENTS

We are grateful to G. D'Angelo, T. Lee, F. Matsuzaki, S. Russell, D. St Johnston, P. Wappner, and M. Welte for fly stocks and antibodies. Stocks were also obtained from the Bloomington *Drosophila* Stock Center (NIH P40OD018537) and the Kyoto *Drosophila* Genetic Resource and Vienna *Drosophila* RNAi Centres. We acknowledge D. Anastasiou, R. Breckenridge, J. Briscoe, E. Cinnamon, P. Driscoll, E. Lana-Elola, R. Makki, C. Poczek, T.J. Ragan, M. Ralser, A. Sawala, I. Stefana, V. Tixier, J.-P. Vincent, and M. Yuneva for advice and critical reading. A.P.B., E.H., J.I.M., and A.P.G. are supported by The Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust, and previously at the National Institute for Medical Research by the UK Medical Research Council (U117584237). A.P.B. was also funded by an EMBO ASTF (194-2012), C.P.L. and C.G. by NIH/NIBIB (P41EB001974) and the Ellison Medical Foundation (AG-SS-2215-08), and A.D.P. by HEFCE and G.K. by the NIHR Respiratory Biomedical Research Unit, University Hospitals, Southampton.

Received: December 18, 2014

Revised: June 15, 2015

Accepted: August 12, 2015

Published: October 8, 2015

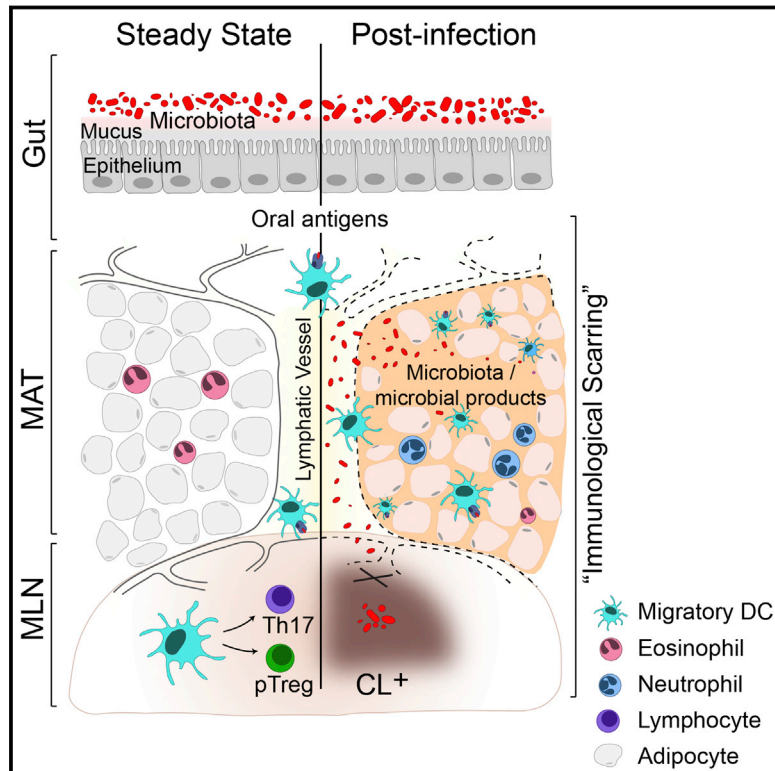
REFERENCES

- Anand, P., Cermelli, S., Li, Z., Kassan, A., Bosch, M., Sigua, R., Huang, L., Ouellette, A.J., Pol, A., Welte, M.A., and Gross, S.P. (2012). A novel role for lipid droplets in the organismal antibacterial response. *eLife* 1, e00003.
- Atkuri, K.R., Mantovani, J.J., Herzenberg, L.A., and Herzenberg, L.A. (2007). N-Acetylcysteine—a safe antidote for cysteine/glutathione deficiency. *Curr. Opin. Pharmacol.* 7, 355–359.
- Bakker, K. (1959). Feeding period, growth and pupation in larvae of *Drosophila melanogaster*. *Entomol. Exp. Appl.* 2, 171–186.
- Barker, D.J. (1995). The Wellcome Foundation Lecture, 1994. The fetal origins of adult disease. *Proc. Biol. Sci.* 262, 37–43.
- Beadle, G.W., Tatum, E.L., and Clancy, C.W. (1938). Food level in relation to rate of development and eye pigmentation in *Drosophila*. *Biol. Bull.* 75, 447–462.
- Bensaad, K., Favaro, E., Lewis, C.A., Peck, B., Lord, S., Collins, J.M., Pinnick, K.E., Wigfield, S., Buffa, F.M., Li, J.L., et al. (2014). Fatty acid uptake and lipid storage induced by HIF-1 α contribute to cell growth and survival under hypoxia-reoxygenation. *Cell Rep.* 9, 349–365.
- Bickel, P.E., Tansey, J.T., and Welte, M.A. (2009). PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. *Biochim. Biophys. Acta* 1791, 419–440.
- Blanpain, C., and Fuchs, E. (2014). Stem cell plasticity. Plasticity of epithelial stem cells in tissue regeneration. *Science* 344, 1242281.
- Bleier, L., and Dröse, S. (2013). Superoxide generation by complex III: from mechanistic rationales to functional consequences. *Biochim. Biophys. Acta* 1827, 1320–1331.
- Britton, J.S., and Edgar, B.A. (1998). Environmental control of the cell cycle in *Drosophila*: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development* 125, 2149–2158.
- Burgess, R.J., Agathocleous, M., and Morrison, S.J. (2014). Metabolic regulation of stem cell function. *J. Intern. Med.* 276, 12–24.
- Carvalho, M., Sampaio, J.L., Palm, W., Brankatschk, M., Eaton, S., and Shevchenko, A. (2012). Effects of diet and development on the *Drosophila* lipidome. *Mol. Syst. Biol.* 8, 600.
- Chandel, N.S., Maltepe, E., Goldwasser, E., Mathieu, C.E., Simon, M.C., and Schumacker, P.T. (1998). Mitochondrial reactive oxygen species trigger hypoxia-induced transcription. *Proc. Natl. Acad. Sci. USA* 95, 11715–11720.
- Chell, J.M., and Brand, A.H. (2010). Nutrition-responsive glia control exit of neural stem cells from quiescence. *Cell* 143, 1161–1173.
- Cheng, L.Y., Bailey, A.P., Leivers, S.J., Ragan, T.J., Driscoll, P.C., and Gould, A.P. (2011). Anaplastic lymphoma kinase spares organ growth during nutrient restriction in *Drosophila*. *Cell* 146, 435–447.
- Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J., and Léopold, P. (2003). A nutrient sensor mechanism controls *Drosophila* growth. *Cell* 114, 739–749.
- Cunningham, L.A., Candelario, K., and Li, L. (2012). Roles for HIF-1 α in neural stem cell function and the regenerative response to stroke. *Behav. Brain Res.* 227, 410–417.
- de la Cova, C., Senoo-Matsuda, N., Ziosi, M., Wu, D.C., Bellosta, P., Quinzii, C.M., and Johnston, L.A. (2014). Supercompetitor status of *Drosophila* Myc cells requires p53 as a fitness sensor to reprogram metabolism and promote viability. *Cell Metab.* 19, 470–483.
- Dumstrei, K., Wang, F., and Hartenstein, V. (2003). Role of DE-cadherin in neuroblast proliferation, neural morphogenesis, and axon tract formation in *Drosophila* larval brain development. *J. Neurosci.* 23, 3325–3335.
- Dunwoodie, S.L. (2009). The role of hypoxia in development of the Mammalian embryo. *Dev. Cell* 17, 755–773.
- Fauny, J.D., Silber, J., and Zider, A. (2005). *Drosophila* Lipid Storage Droplet 2 gene (Lsd-2) is expressed and controls lipid storage in wing imaginal discs. *Dev. Dyn.* 232, 725–732.
- Grant, J., Saldanha, J.W., and Gould, A.P. (2010). A *Drosophila* model for primary coenzyme Q deficiency and dietary rescue in the developing nervous system. *Dis. Model. Mech.* 3, 799–806.
- Gruenewald, P. (1963). Chronic Fetal Distress and Placental Insufficiency. *Biol. Neonat.* 5, 215–265.
- Gutierrez, E., Wiggins, D., Fielding, B., and Gould, A.P. (2007). Specialized hepatocyte-like cells regulate *Drosophila* lipid metabolism. *Nature* 445, 275–280.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Homem, C.C., and Knoblich, J.A. (2012). *Drosophila* neuroblasts: a model for stem cell biology. *Development* 139, 4297–4310.
- Homem, C.C., Steinmann, V., Burkard, T.R., Jais, A., Esterbauer, H., and Knoblich, J.A. (2014). Ecdysone and mediator change energy metabolism to terminate proliferation in *Drosophila* neural stem cells. *Cell* 158, 874–888.
- Kamphorst, J.J., Cross, J.R., Fan, J., de Stanchina, E., Mathew, R., White, E.P., Thompson, C.B., and Rabinowitz, J.D. (2013). Hypoxic and Ras-transformed cells support growth by scavenging unsaturated fatty acids from lysophospholipids. *Proc. Natl. Acad. Sci. USA* 110, 8882–8887.
- Kis, V., Barti, B., Lippai, M., and Sass, M. (2015). Specialized Cortex Glial Cells Accumulate Lipid Droplets in *Drosophila* melanogaster. *PLoS ONE* 10, e0131250.
- Kühnlein, R.P. (2011). The contribution of the *Drosophila* model to lipid droplet research. *Prog. Lipid Res.* 50, 348–356.

- Kühnlein, R.P. (2012). Thematic review series: Lipid droplet synthesis and metabolism: from yeast to man. Lipid droplet-based storage fat metabolism in *Drosophila*. *J. Lipid Res.* 53, 1430–1436.
- Lanet, E., Gould, A.P., and Maurange, C. (2013). Protection of neuronal diversity at the expense of neuronal numbers during nutrient restriction in the *Drosophila* visual system. *Cell Rep.* 3, 587–594.
- Lee, K.E., and Simon, M.C. (2012). From stem cells to cancer stem cells: HIF takes the stage. *Curr. Opin. Cell Biol.* 24, 232–235.
- Li, X., Chen, Z., and Desplan, C. (2013). Temporal patterning of neural progenitors in *Drosophila*. *Curr. Top. Dev. Biol.* 105, 69–96.
- Liu, L., Zhang, K., Sandoval, H., Yamamoto, S., Jaiswal, M., Sanz, E., Li, Z., Hui, J., Graham, B.H., Quintana, A., and Bellen, H.J. (2015). Glial lipid droplets and ROS induced by mitochondrial defects promote neurodegeneration. *Cell* 160, 177–190.
- Maurange, C., Cheng, L., and Gould, A.P. (2008). Temporal transcription factors and their targets schedule the end of neural proliferation in *Drosophila*. *Cell* 133, 891–902.
- Mohyeldin, A., Garzón-Muvdi, T., and Quiñones-Hinojosa, A. (2010). Oxygen in stem cell biology: a critical component of the stem cell niche. *Cell Stem Cell* 7, 150–161.
- Nakaya, K., Ushio, H., Matsukawa, S., Shimizu, M., and Ohshima, T. (2005). Effects of droplet size on the oxidative stability of oil-in-water emulsions. *Lipids* 40, 501–507.
- Negre-Salvayre, A., Coatrieux, C., Ingueneau, C., and Salvayre, R. (2008). Advanced lipid peroxidation end products in oxidative damage to proteins. Potential role in diseases and therapeutic prospects for the inhibitors. *Br. J. Pharmacol.* 153, 6–20.
- Okano, H., and Temple, S. (2009). Cell types to order: temporal specification of CNS stem cells. *Curr. Opin. Neurobiol.* 19, 112–119.
- Olsen, L.F., Issinger, O.G., and Guerra, B. (2013). The Yin and Yang of redox regulation. *Redox Rep.* 18, 245–252.
- Parra-Peralbo, E., and Culi, J. (2011). *Drosophila* lipophorin receptors mediate the uptake of neutral lipids in oocytes and imaginal disc cells by an endocytosis-independent mechanism. *PLoS Genet.* 7, e1001297.
- Pearson, B.J., and Doe, C.Q. (2004). Specification of temporal identity in the developing nervous system. *Annu. Rev. Cell Dev. Biol.* 20, 619–647.
- Pugh, C.W., and Ratcliffe, P.J. (2003). Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* 9, 677–684.
- Rapport, E.W., Stanley-Samuelson, D., and Dadd, R.H. (1983). Ten generations of *Drosophila melanogaster* reared axenically on a fatty acid-free holidic diet. *Arch. Insect Biochem. Physiol.* 1, 243–250.
- Rysman, E., Brusselmans, K., Scheys, K., Timmermans, L., Derua, R., Munck, S., Van Veldhoven, P.P., Waltregny, D., Daniëls, V.W., Machiels, J., et al. (2010). De novo lipogenesis protects cancer cells from free radicals and chemotherapeutics by promoting membrane lipid saturation. *Cancer Res.* 70, 8117–8126.
- Scadden, D.T. (2006). The stem-cell niche as an entity of action. *Nature* 441, 1075–1079.
- Skeath, J.B., and Thor, S. (2003). Genetic control of *Drosophila* nerve cord development. *Curr. Opin. Neurobiol.* 13, 8–15.
- Slack, J.M. (2008). Origin of stem cells in organogenesis. *Science* 322, 1498–1501.
- Sousa-Nunes, R., Cheng, L.Y., and Gould, A.P. (2010). Regulating neural proliferation in the *Drosophila* CNS. *Curr. Opin. Neurobiol.* 20, 50–57.
- Sousa-Nunes, R., Yee, L.L., and Gould, A.P. (2011). Fat cells reactivate quiescent neuroblasts via TOR and glial insulin relays in *Drosophila*. *Nature* 471, 508–512.
- Spéder, P., and Brand, A.H. (2014). Gap junction proteins in the blood-brain barrier control nutrient-dependent reactivation of *Drosophila* neural stem cells. *Dev. Cell* 30, 309–321.
- Steinhauser, M.L., Bailey, A.P., Senyo, S.E., Guillemer, C., Perlstein, T.S., Gould, A.P., Lee, R.T., and Lechene, C.P. (2012). Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism. *Nature* 481, 516–519.
- Taverna, E., Götz, M., and Huttner, W.B. (2014). The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell Dev. Biol.* 30, 465–502.
- Tennessen, J.M., Baker, K.D., Lam, G., Evans, J., and Thummel, C.S. (2011). The *Drosophila* estrogen-related receptor directs a metabolic switch that supports developmental growth. *Cell Metab.* 13, 139–148.
- Tower, J. (2012). Stress and stem cells. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 789–802.
- Truman, J.W., and Bate, M. (1988). Spatial and temporal patterns of neurogenesis in the central nervous system of *Drosophila melanogaster*. *Dev. Biol.* 125, 145–157.
- Uchida, K. (2003). 4-Hydroxy-2-nonenal: a product and mediator of oxidative stress. *Prog. Lipid Res.* 42, 318–343.
- Ushio-Fukai, M., and Rehman, J. (2014). Redox and metabolic regulation of stem/progenitor cells and their niche. *Antioxid. Redox Signal.* 21, 1587–1590.
- Walther, T.C., and Farese, R.V., Jr. (2012). Lipid droplets and cellular lipid metabolism. *Annu. Rev. Biochem.* 81, 687–714.
- Wang, K., Zhang, T., Dong, Q., Nice, E.C., Huang, C., and Wei, Y. (2013). Redox homeostasis: the linchpin in stem cell self-renewal and differentiation. *Cell Death Dis.* 4, e537.
- Wilfling, F., Wang, H., Haas, J.T., Krahmer, N., Gould, T.J., Uchida, A., Cheng, J.X., Graham, M., Christiano, R., Fröhlich, F., et al. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocating from the ER to lipid droplets. *Dev. Cell* 24, 384–399.
- Wilfling, F., Haas, J.T., Walther, T.C., and Farese, R.V., Jr. (2014). Lipid droplet biogenesis. *Curr. Opin. Cell Biol.* 29, 39–45.
- Wingrove, J.A., and O'Farrell, P.H. (1999). Nitric oxide contributes to behavioral, cellular, and developmental responses to low oxygen in *Drosophila*. *Cell* 98, 105–114.
- Young, S.G., and Zechner, R. (2013). Biochemistry and pathophysiology of intravascular and intracellular lipolysis. *Genes Dev.* 27, 459–484.

Microbiota-Dependent Sequelae of Acute Infection Compromise Tissue-Specific Immunity

Graphical Abstract



Authors

Denise Morais da Fonseca,
Timothy W. Hand, Seong-Ji Han, ...,
Ronald N. Germain,
Gwendalyn J. Randolph, Yasmine Belkaid

Correspondence

ybelkaid@niaid.nih.gov

In Brief

Following a single episode of acute gastrointestinal infection, microbiota-dependent inflammation of the mesenteric adipose tissue and damage to the lymphatic system disrupt the communication between the immune system and the tissue, causing an "immunological scar" that compromises tissue immunity and homeostasis in the long term.

Highlights

- Acute infections can persistently compromise tissue-specific immunity
- Acute infections induce immune remodeling of the adipose tissue
- Infection-induced lymphatic leakage deviates migratory DC trafficking
- Microbiota sustains mesenteric inflammation and immune dysfunction post-infection



Microbiota-Dependent Sequelae of Acute Infection Compromise Tissue-Specific Immunity

Denise Morais da Fonseca,^{1,2,3,12} Timothy W. Hand,^{1,2,12,13} Seong-Ji Han,^{1,2} Michael Y. Gerner,⁴ Arielle Glatman Zaretsky,^{1,2} Allyson L. Byrd,^{1,2,5,6} Oliver J. Harrison,^{1,2} Alexandra M. Ortiz,⁷ Mariam Quinones,⁸ Giorgio Trinchieri,⁹ Jason M. Brechley,⁷ Igor E. Brodsky,¹⁰ Ronald N. Germain,⁴ Gwendalyn J. Randolph,¹¹ and Yasmine Belkaid^{1,2,*}

¹Mucosal Immunology Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIAID/NIH), Bethesda, MD 20892, USA

²Immunity at Barrier Sites Initiative, NIAID/NIH, Bethesda, MD 20892, USA

³Department of Biochemistry and Immunology, Ribeirão Preto School of Medicine, University of São Paulo, Ribeirão Preto, 14049-900, Brazil

⁴Lymphocyte Biology Section, Laboratory of Systems Biology, NIAID/NIH, Bethesda, MD 20892, USA

⁵Translational and Functional Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA

⁶Department of Bioinformatics, Boston University, Boston, MA 02215, USA

⁷Program in Tissue Immunity and Repair and Immunopathogenesis Section, Laboratory of Molecular Microbiology, NIAID, NIH, Bethesda, MD 20892, USA

⁸Bioinformatics and Computational Bioscience Branch, NIAID/NIH, Bethesda, MD 20892, USA

⁹Cancer and Inflammation Program, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA

¹⁰Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

¹¹Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

¹²Co-first author

¹³Present address: Richard King Mellon Foundation Institute for Pediatric Research, Department of Pediatrics, University of Pittsburgh, Pittsburgh, PA 15224, USA

*Correspondence: ybelkaid@niaid.nih.gov

<http://dx.doi.org/10.1016/j.cell.2015.08.030>

SUMMARY

Infections have been proposed as initiating factors for inflammatory disorders; however, identifying associations between defined infectious agents and the initiation of chronic disease has remained elusive. Here, we report that a single acute infection can have dramatic and long-term consequences for tissue-specific immunity. Following clearance of *Yersinia pseudotuberculosis*, sustained inflammation and associated lymphatic leakage in the mesenteric adipose tissue deviates migratory dendritic cells to the adipose compartment, thereby preventing their accumulation in the mesenteric lymph node. As a consequence, canonical mucosal immune functions, including tolerance and protective immunity, are persistently compromised. Post-resolution of infection, signals derived from the microbiota maintain inflammatory mesentery remodeling and consequently, transient ablation of the microbiota restores mucosal immunity. Our results indicate that persistent disruption of communication between tissues and the immune system following clearance of an acute infection represents an inflection point beyond which tissue homeostasis and immunity is compromised for the long-term.

INTRODUCTION

The immune system deals with highly diverse infectious challenges in a manner that both promotes the control of invading agents and restores tissue homeostasis. At barrier tissues, the sites of constitutive microbial exposure, tissue-specific immunity requires defined structural components that provide spatial segregation from the microbiota to ensure maintenance of organ function and enable the balance between tolerance to environmental antigens and regulated protective immunity (Mowat and Agace, 2014). These balanced processes are essential for the preservation of tissue integrity. Inflammatory diseases affecting barrier sites develop as a result of a breakdown in tissue homeostasis and the subsequent failure to regulate immune responses to environmental or microbial antigens (Belkaid and Segre, 2014; Maloy and Powrie, 2011).

Acute infections represent frequent and violent tissue perturbations from which the immune system must rapidly rebound. It is estimated that >70% of all people in the United States experience a respiratory tract infection each year, and the average child will suffer ten diarrheal episodes before the age of 5, a number that is dramatically higher in low to middle income countries (Fendrick et al., 2003; Kosek et al., 2003; Vernacchio et al., 2006). At barrier sites in particular, acute infections represent highly volatile situations, with pathogens, commensals, and environmental antigens transiently sharing the same inflamed environment. As a consequence, in the gastrointestinal (GI) tract, acute mucosal infections are characterized by dysbiosis associated with significant shifts in the microbiota and dominance of

bacteria with enhanced invasive capabilities and other properties that can directly exacerbate inflammation and tissue damage (Belkaid and Hand, 2014). Because regulatory responses are often transiently neutralized during acute infection, these encounters can also blur the distinction between benign constituents and pathogenic organisms (DePaolo et al., 2011; Hand et al., 2012; Oldenhove et al., 2009). What remains unclear is whether acute infections produce persistent changes that negatively impact immune function, and if so, what are the mechanisms of such effects?

These questions are especially compelling in light of the rising incidence of inflammatory disorders, particularly involving those diseases affecting barrier tissues such as inflammatory bowel disease, psoriasis, allergy, and asthma (Molodecky et al., 2012; Salgame et al., 2013). A number of environmental stressors, including infections have been invoked as possible triggers. Chronic infections in particular can have significant bystander consequences on the capacity of the host to respond to subsequent challenges (Barton et al., 2007; Cadwell et al., 2010; Jamieson et al., 2013; Osborne et al., 2014; Salgame et al., 2013). However, identifying direct associations between defined infectious agents and the initiation of chronic disease has remained difficult. One possible explanation for this difficulty in establishing cause-effect relationships could be that infections associated with the breakdown of tissue homeostasis may not be temporally coincident with the eruption of symptomatic disease, and previously cleared infections could have long-term and cumulative effects on the immune system.

Here, we show that a single transient encounter with a GI pathogen, *Yersinia pseudotuberculosis*, can have dramatic and persistent consequences on tissue-specific immunity. Infection-induced remodeling of the mesentery and increased leakage of lymphatic vessels persistently interrupts the immune dialog between the gut and associated lymphoid tissue. These sequelae of a cured infection interfere with the capacity of the host to develop both tolerance and adaptive immunity to oral antigens and are sustained by the microbiota. Thus, in defined infectious settings, the immune system can reach an inflection point beyond which restoration of tissue homeostasis and immunity requires intervention. These results reveal that acute infections could impose “immunological scarring” that may have cumulative and long-term consequences. This work provides a framework for understanding how previously experienced infectious stressors that are no longer present in the host can cause localized immune damage that may contribute to the breakdown of tissue-specific immunity and the emergence of complex diseases.

RESULTS

Oral Infection with *Yersinia pseudotuberculosis* Induces Chronic Mesenteric Lymphadenopathy

Yersinia pseudotuberculosis is a foodborne Gram-negative bacterium that, in humans, can cause a range of GI syndromes from acute enteritis to mesenteric lymphadenitis and pseudoappendicitis (Asano, 2012; Wren, 2003). In mice, oral infection with 10^7 colony-forming units (CFU) of *Y. pseudotuberculosis* (IP32777) caused a transient infection, with the peak of bacterial burden

at day 7 post-infection (Figure 1A) (McPhee et al., 2012; Simonet and Falkow, 1992). As previously described (Fahlgren et al., 2014), systemic control of the infection was achieved by 3 weeks post-infection in most animals (Figures 1A and S1A). T cell responses are important for the control of *Y. pseudotuberculosis* infection (Bergman et al., 2009; Bergsbaken and Bevan, 2015; Zhang et al., 2012) and the number of IFN- γ -producing antigen-specific (YopE_{69–77}) CD8⁺ T cells increased significantly by 2 weeks post-infection before contracting, in concert with clearance of the bacteria (Figures S1B and S1C).

Concurrent with the peak of bacterial burden, *Y. pseudotuberculosis* induced a rapid influx of neutrophils and monocytes into all infected tissues (Figures 1B–1D and data not shown). This increase in phagocytic cells largely subsides in all compartments, except for the MLN (Figures 1B–1D and data not shown). Despite efficient control of the bacteria, a significant fraction of infected mice (70%) developed chronic mesenteric lymphadenopathy (CL⁺) defined by 3- to 4-fold tissue enlargement, the formation of central abscesses, the presence of foamy macrophages, and significant collagen deposition (Figures 1E–1H and S1D). All of these features are reminiscent of the MLN pathology that can be observed during human infections with *Y. pseudotuberculosis* (Asano, 2012). Lymphadenopathy was highly restricted to the MLN and still detectable 9 months post-resolution of the infection (Figures 1E and 1I). This response was not associated with persistent infection of the gut, MLN, and mesenteric adipose tissue (MAT) (Figures 1A and S1A). However, CL⁺ lymph nodes were not sterile and bacteria (the vast majority belonging to the genus *Lactobacillus*) could be grown post-resolution of the infection (Figures S1E and S1F). Thus, acute infection with *Y. pseudotuberculosis* can induce long-term and localized damage to gut-associated lymphoid structures.

Impaired Tissue-Specific Adaptive Immunity Post *Y. pseudotuberculosis* Infection

The profound disruption of MLN structure following *Y. pseudotuberculosis* infection led us to explore the possibility that the homeostatic immune dialog between mucosal antigens and the gut-associated secondary lymphoid tissue may be compromised. To address this point, we first explored the potential impact of lymphoid tissue remodeling following infection on the acquisition of oral tolerance, a process that is largely regulated by the induction of antigen-specific peripheral T_{reg} cells (pT_{reg}) (Coombes et al., 2007; Hadis et al., 2011; Sun et al., 2007). To this end, T cells from *Rag1*^{−/−} OT-II TCR-transgenic mice were transferred into mice fed with ovalbumin (OVA). Controls included naive hosts as well as mice that controlled *Y. pseudotuberculosis* without developing lymphadenopathy (CL[−]). In naive OVA-fed mice, a significant fraction of OT-II cells accumulating in the GALT expressed Foxp3 (Figures 2A and 2B). Furthermore, >30% of pT_{reg} cells co-expressed GATA3 (Figure 2A), a transcription factor associated with T_{reg} fitness (Wang et al., 2011; Wohlfert et al., 2011). In contrast, pT_{reg} generation was significantly impaired in mice harboring enlarged MLNs (CL⁺) compared to naive or CL[−] mice (Figures 2A and 2B). The percentage of T_{reg} cells expressing GATA3 was also significantly decreased in CL⁺ compared to control mice (Figures 2A and 2B). Acquisition of tolerance to orally fed antigens can be

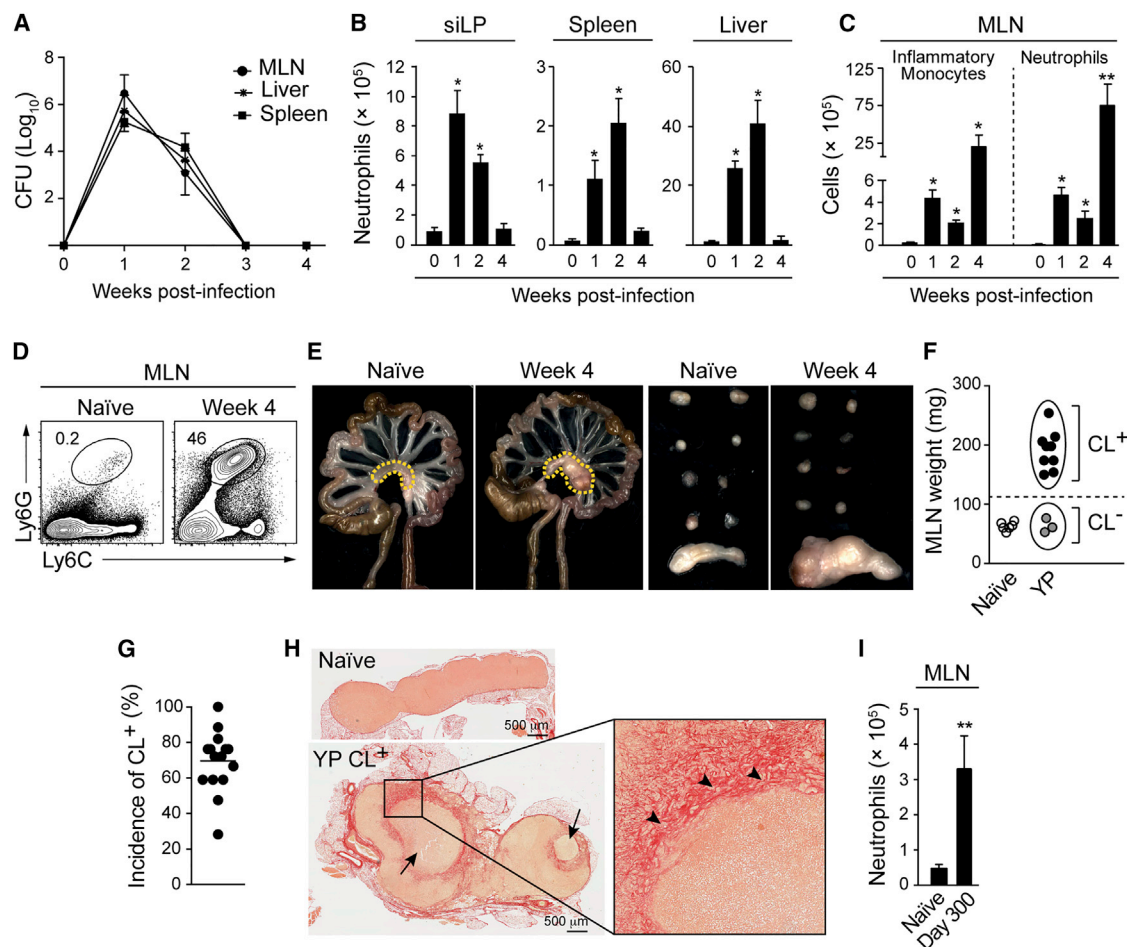


Figure 1. Oral Infection with *Y. pseudotuberculosis* Induces Persistent Mesenteric Lymphadenopathy

C57BL/6 mice were orally infected with *Y. pseudotuberculosis* (YP).

(A) Infectious burden in the spleen, MLN, and liver at indicated time points.

(B and C) Numbers of neutrophils and inflammatory monocytes from siLP, spleen, liver, and MLN.

(D) Representative flow cytometric contour plots indicating the percentage of neutrophils in the MLNs of naïve and infected mice.

(E) Images on the left depict the gastrointestinal tract, mesenteric adipose tissue (MAT), and MLN (dotted lines) from naïve and infected mice. Images on the right show (from top) the axillary, brachial, lumbar, inguinal, and mesenteric lymph nodes.

(F) Weight of MLN from naïve (white circles) or infected mice with chronic lymphadenopathy (CL⁺) (black circles) and without lymphadenopathy (CL⁻) (gray circles).

(G) Compilation of 15 separate experiments involving week 4 to week 9 infected mice showing the percent of animals with lymphadenopathy.

(H) Histology of naïve and CL⁺ MLNs stained with picosirius red. Large arrows show abscesses, arrowheads indicate areas of collagen deposition.

(I) Number of neutrophils in the MLN at day 300 post-infection. All bar graphs show the mean ± SEM. All data shown, except (I) and (G), is representative of three to six experiments, each containing three to five naïve and six to ten infected animals. Data represented in (I) are representative of 2 experiments with three naïve and three to nine infected animals. **p* < 0.05, ***p* < 0.005 compared to naïve mice (Student's *t* test).

See also Figure S1.

measured by assessing immune responses following challenge at peripheral sites (Curotto de Lafaille et al., 2008; Weiner et al., 2011; Worbs et al., 2006). In contrast to naïve mice, CL⁺ mice fed with OVA failed to acquire oral tolerance and developed a delayed type hypersensitivity response to OVA challenge (Figure 2C). Thus, mesenteric lymphadenopathy induced by infection is associated with impaired acquisition of oral tolerance to food antigens.

We next assessed whether lymphadenopathy could potentially interfere with the induction of mucosal effector responses.

To this end, we utilized a model of oral vaccination combining a mixture of OVA and an attenuated double mutant of the heat-labile enterotoxin of enterotoxigenic *Escherichia coli* (LT R192G/L211A; hereafter referred to as dmLT) (Hall et al., 2008). As previously shown, this regimen induces robust OVA and dmLT-specific Th17 cells (Hall et al., 2008) (Figures 2D and 2E). The induction of Th17 cells specific for dmLT or OVA was dramatically impaired in CL⁺ mice compared to controls (Figures 2D and 2E). Induction of dmLT-specific and OVA-specific IgA was also significantly reduced in CL⁺ mice

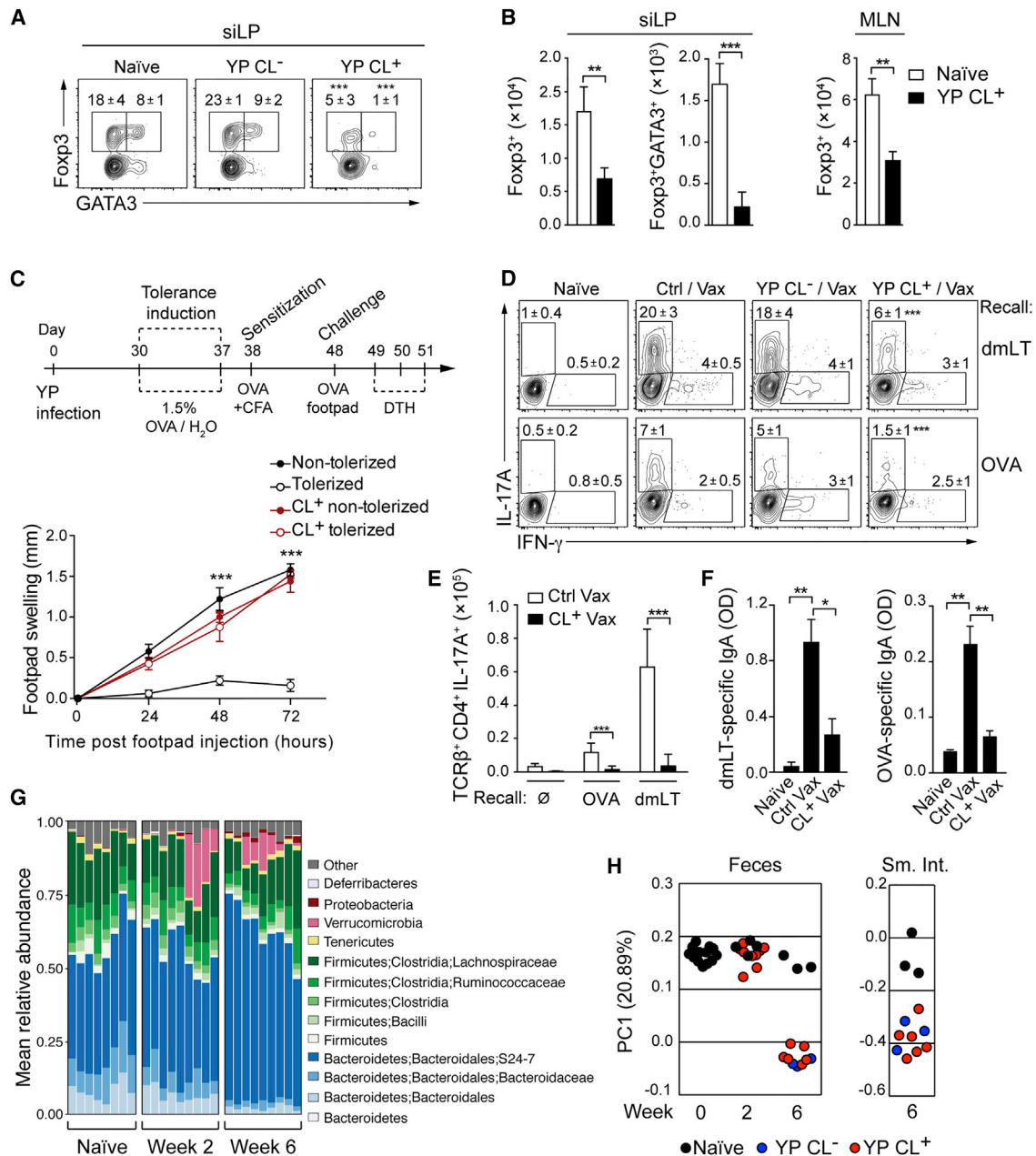


Figure 2. Infection-Induced Mesenteric Lymphadenopathy Is Associated with Disruption of Mucosal Immunity

(A and B) Naïve or 4-week infected C57BL/6 (CD45.2) mice were transferred with CD45.1⁺ *Rag1*^{-/-} OT-II TCR transgenic T cells (CD45.1), fed ovalbumin (OVA) in the drinking water, and cells from the siLP and MLNs were isolated. (A) Representative flow cytometric contour plots of Foxp3 and GATA3 expression by OT-II T cells isolated from the siLP of naïve and infected mice with or without lymphadenopathy (YP CL⁺ and YP CL⁻, respectively). Numbers represent the mean percentage within the gate (±SEM) of all samples in this experiment. (B) Numbers of OTII Foxp3⁺ and OT-II Foxp3⁺GATA3⁺ T cells in the siLP and MLN.

(C) Top: scheme for induction of oral tolerance. CFA, complete Freund's adjuvant. Bottom: footpad swelling was measured in the feet of sensitized mice after challenge with OVA.

(D–F) Naïve, YP CL⁻ and YP CL⁺ mice were immunized orally with OVA and double mutant heat labile toxin (dmLT). (D and E) Seven days after immunization, lymphocytes were isolated from the siLP and stimulated in vitro with DCs pulsed with dmLT or OVA to measure T cell responses. (D) Flow cytometric contour plots show representative populations of antigen-specific IFN-γ and IL-17A-producing CD4⁺ T cells from naïve, vaccinated (Ctrl/Vax), or *Y. pseudotuberculosis*-infected vaccinated (YP CL^{+/}-Vax) mice. Numbers in plots represent the mean frequency of cells within the adjacent gate (±SEM) of all samples within this experiment. (E) Numbers of IL-17A-producing CD4⁺ T cells from (D). (F) Measurement of dmLT and OVA-specific fecal IgA by ELISA. OD, optical density.

(legend continued on next page)

compared to vaccinated controls (Figure 2F). Infection with *Y. pseudotuberculosis* also modestly modifies the fecal and small intestinal composition of the microbiota (Figures 2G and 2H). Of note, no differences were found between CL⁺ versus CL⁻ mice (Figure 2H).

Together, these results reveal that a single acute infection can chronically alter the ability of the immune system to develop canonical mucosal responses necessary for maintenance of gut immune homeostasis.

Chronic Mesenteric Lymphadenopathy Is Associated with Defects in Dendritic Cell Migration

Effective activity of the mucosal immune system relies upon efficient trafficking of DCs to the MLN to ensure the maintenance of homeostasis with respect to antigens acquired from the intestinal lumen (Hooper and Macpherson, 2010). We first examined the spatial organization of the MLN following infection with *Y. pseudotuberculosis*. Consistent with flow cytometry analysis (Figures 1B–1D), confocal imaging revealed a significant accumulation of LysM-eGFP⁺ cells at the center of the MLN (Figure 3A). The central T cell zone was displaced and distributed in a disorganized manner at the periphery of the MLN of CL⁺ (Figures 3A and S2A). Additionally, B cell follicles and lymphatic vessels that, in naive mice, localize at the periphery of the MLN, were located at the edge of the abscess (Figure 3A). Naive T cells were still present in the MLN post-infection, suggesting that some MLN functions, such as the trafficking of lymphocytes from the blood, may have been preserved (Figure S2B).

In the gut, immune dialog is mediated by three main antigen presenting cells that can be segregated based on their expression of integrin α E (CD103) and α M (CD11b) (Figure S2C) (Bekiaris et al., 2014; Bogunovic et al., 2009; Grainger et al., 2014). In particular, the migratory CD103-expressing DC subsets have been associated with the induction of Th17 and T_{reg} cells (Coombes et al., 2007; Satpathy et al., 2013; Schlitzner et al., 2013; Schulz et al., 2009; Sun et al., 2007; Varol et al., 2009). The frequency and absolute number of migratory CD103⁺CD11b⁺ DCs were dramatically reduced in the MLN of CL⁺ mice compared to naive and CL⁻ mice (Figures 3B and 3C). CD103⁺CD11b⁻ DCs in the MLN were also reduced, albeit to a lesser degree (Figures 3B and 3C). For the remainder of this study, we focused on the migratory CD103⁺CD11b⁺ DC subset, whose unique phenotype allows for faithful tracking during inflammation. Importantly, the defect in CD103⁺CD11b⁺ DCs was restricted to the MLN as gut lamina propria DC subsets were unchanged in CL⁺ mice versus controls (Figure 3D). These results support the idea that lack of migratory CD103⁺CD11b⁺ DCs in the MLN was not the consequence of a defect in DC development or gut homing, but the result of impaired migration and/or accumulation in the MLN.

Extraction of cells from the MLNs of CL⁺ mice could be biased because of the high level of collagen deposition (Figure 1H). Histo-cytometry allows for the identification, spatial positioning,

and quantitative measurement of cells within their host tissue (Gerner et al., 2012). Because CD103 is highly sensitive to fixation, we devised an alternative gating strategy to identify CD103⁺CD11b⁺ DCs by histo-cytometry (Figures S3A and S3B) (Merad et al., 2013). As previously described (Gerner et al., 2012; Kissenpfennig et al., 2005), migratory and resident DC subsets localized in distinct regions of naive LNs (Figures 3E, 3G, and S3C). This approach allowed us to describe the localization of the migratory CD103⁺CD11b⁺ DCs in the MLN under steady-state conditions as positioned peripheral to CD103⁺CD11b⁻ DCs on the edge of the putative T cell zone and in close proximity to the B cell follicles (Figures 3E and 3G). Following infection, the CD103⁺CD11b⁺ DCs were virtually absent from CL⁺ MLNs compared to controls (Figures 3E–3G, S3C, and S3D). The distribution of other DC subsets was not altered (Figures 3E, S3C, and S3D). Thus, two complementary approaches revealed a sustained defect in the accumulation of migratory DCs in the MLN post-infection.

Acute Infection Alters Lymphatic Integrity and Mesenteric Adipose Tissue Homeostasis

Notwithstanding the direct infection-related structural disruption of the MLN, our results thus far support the idea that impaired mucosal immunity following infection may be a consequence of a selective defect in the capacity of migratory DCs to access the MLN. In addition to cell trafficking, the gut-associated lymphatic system is also responsible for the transport of lipids. We assessed the physical integrity of gut-draining lymphatic vessels by orally administering the fluorescent long-chain fatty acid, Bodipy FL C16 (Khalifeh-Soltani et al., 2014; Randolph and Miller, 2014). In contrast to control mice, Bodipy FL C16 leaked into the surrounding mesenteric adipose tissue (MAT) of CL⁺ mice, revealing a major defect in lymphatic containment, a defect that was sustained up to 10 months post-infection (Figures 4A and S4A).

We next assessed how the loss of lymphatic integrity could impact MAT immune homeostasis. Imaging of the MAT revealed a massive infiltrate of hematopoietic cells in CL⁺, but not CL⁻ mice (Figure 4B). This infiltrate was associated with a significant increase in neutrophils and macrophages while the number of type 2 immune cells including eosinophils and ILC2s, was significantly reduced (Figure 4C). The MAT shifted toward type 1 immunity, with a significant increase in the transcription of inflammatory mediators, including, *Ifng*, *Il1a*, *Il1b*, *Tnfa*, *Il12*, *Fpr1*, *Fpr2*, antimicrobial molecules, and chemokines (Figures 4D and S4B; Table S1). Alteration of lymphatic integrity was not unique to *Y. pseudotuberculosis*, as is demonstrated by leakage in models of T cell transfer colitis and *Toxoplasma gondii* infection (Figures S4C and S4D).

At steady state, the MAT contains CD103⁺CD11b⁻ and CD103⁻CD11b⁺ DCs, while CD103⁺CD11b⁺ DCs are largely absent from this compartment (Figures 5A, 5B, and S4E). In contrast, the frequency and number of CD103⁺CD11b⁺ DCs

(G) Bar graphs showing the fraction of the fecal microbiota represented by individual operational taxonomic units (OTUs) identified by 16S bacterial gene sequencing at the time points indicated.

(H) Principal coordinate analysis of 16S gene sequencing data derived from fecal and small intestinal samples (weighted UniFrac). All experiments, except (G) and (H), are representative of three to five separate experiments containing five control mice and five to ten infected mice per group. All graphs show the mean \pm SEM.

*p < 0.05, **p < 0.005, ***p < 0.0005 (Student's t test).

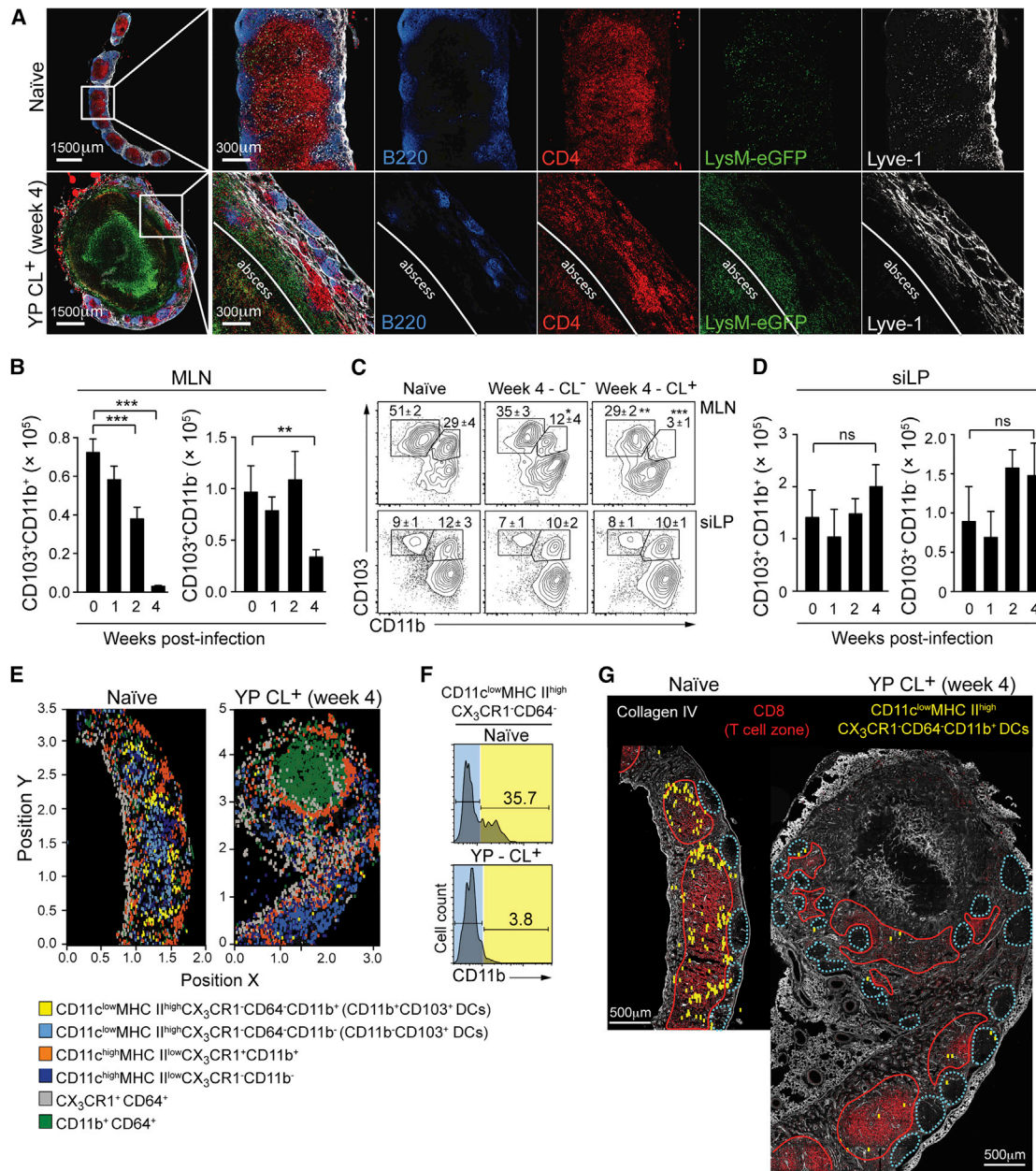


Figure 3. Mesenteric Lymphadenopathy Is Associated with a Reduction in Migratory DCs in the MLN

(A) LysM-eGFP reporter mice were infected with *Y. pseudotuberculosis*. At 4 weeks post-infection, sections of naïve or CL⁺ MLNs were analyzed by confocal microscopy.

(B–D) Bar graphs show the numbers of each DC subset (gated as described in Figure S2C), isolated from the MLN (B) and siLP (D) at the time points indicated. (C) Representative flow cytometric contour plots of DC subsets from the MLN and siLP. Numbers in plot represent the mean frequency of cells within the adjacent gate (±SEM) of all samples within this experiment.

(E–G) MLN sections from naïve or 4-week infected CX₃CR1-GFP reporter mice were stained for analysis by multiparameter confocal microscopy. (E) Three-dimensional “Cell-of-Interest” surfaces were generated based on CD11c expression and quantitatively analyzed by histo-cytometry (Figure S3) to identify the X and Y positions of individual cells within the antigen-presenting cell subsets (labels below image). (F) Representative histograms obtained from the histo-cytometric analysis indicating the expression of CD11b⁺ in cells occupying the CD11c⁺ MHC II^{high} CX₃CR1⁻ CD64⁻ gate (Figure S3). (G) Representative confocal image of MLN sections stained for Collagen IV and CD8α overlaid with the position of CD11c⁺ MHC II^{high} CX₃CR1⁻ CD64⁻ CD11b⁺ DCs (analogous to CD11b⁺CD103⁺ DCs) reconstructed by histo-cytometry. Dotted blue lines represent the B cell zones and red lines indicate the T cell zones. Images in (A), (E), and (G) are representative of two separate experiments. Flow cytometry data from (B)–(D) are representative of five experiments each containing five control animals and five to ten infected animals. All bar graphs show the mean ± (SEM). *p < 0.05, **p < 0.005, ***p < 0.0005 (Student’s t test). CL⁺, infected mice with lymphadenopathy; CL⁻, infected mice without lymphadenopathy.

See also Figures S2 and S3.

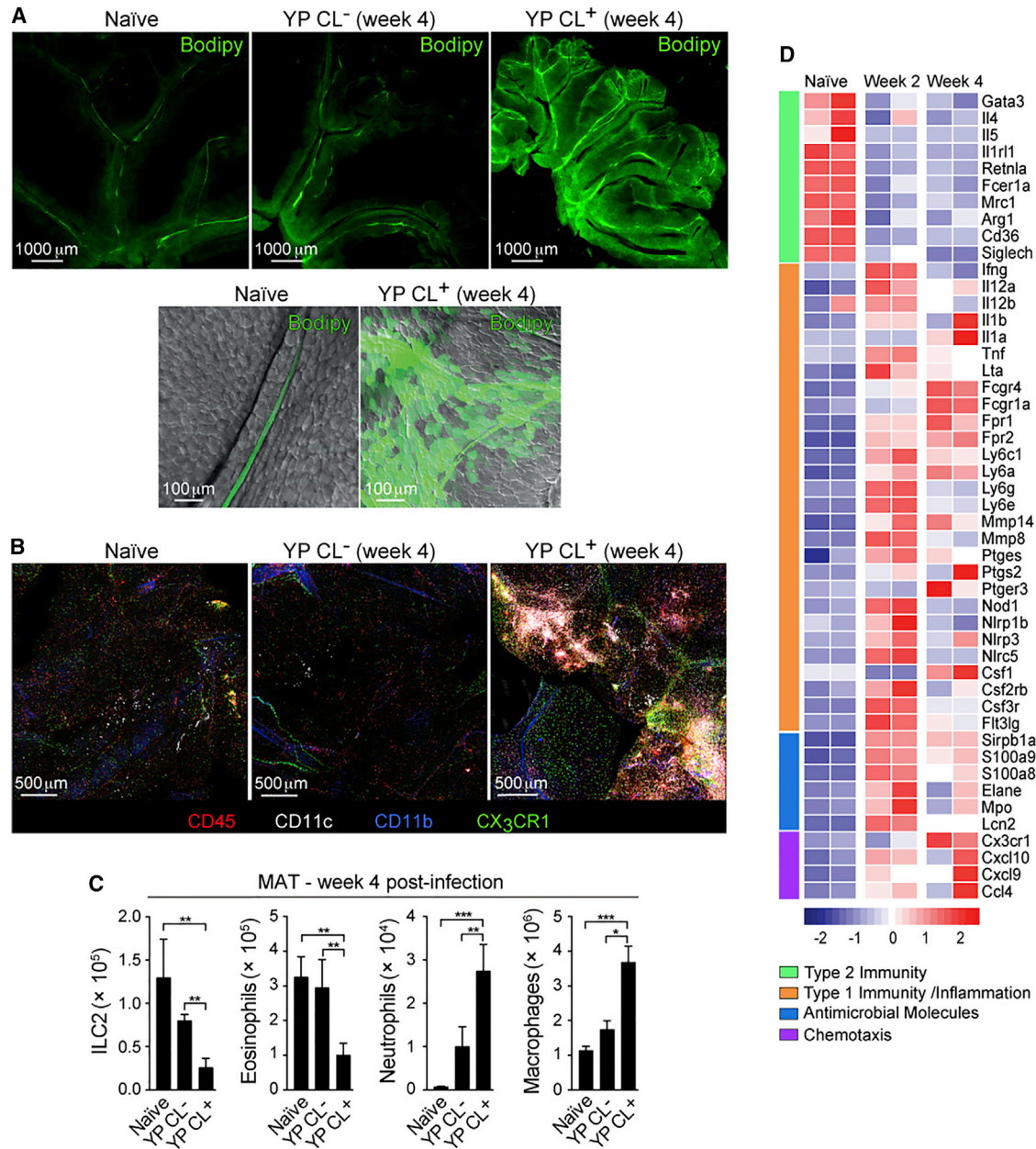


Figure 4. *Yersinia pseudotuberculosis* Infection Is Associated with Lymphatic Leakage and Mesenteric Adipose Tissue Remodeling

(A) Naive or 4-week infected mice were gavaged with Bodipy FL C16. Shown are representative fluorescent microscope images of the MAT (top) or high-magnification confocal images of individual MAT lymphatics overlaid with the bright field (bottom).

(B) MAT samples from naive or infected CX3CR1-GFP mice were stained and imaged by confocal microscopy.

(C) Numbers of ILC2s, eosinophils, neutrophils, and macrophages in the MAT determined by flow cytometry.

(D) Heat map of gene expression from CD45⁺ cells isolated from the MAT at the indicated time points and analyzed by NanoString technology. Images in (A) and (B) are representative of three experiments. Data from (C) are representative of three to five experiments each containing five control animals and five to ten infected animals. Data in (D) are representative of two separate experiments. All bar graphs show the mean (\pm SEM). * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$ (Student's t test). CL⁺, infected mice with lymphadenopathy; CL⁻, infected mice without lymphadenopathy.

See also Figure S4 and Table S1.

dramatically increased in the MAT post-infection, an effect that persisted for at least 10 months post-infection (Figures 5A and 5B). We next examined whether migratory DCs had extravasated into the adipose tissue by confocal microscopy. While, in control

mice few CD11c⁺ and/or CD11b⁺ cells resided in the adipose tissue, in CL⁺ mice, antigen-presenting cells accumulated in the MAT particularly in areas surrounding the lymphatic vessels (Figures 5C, 5D, and S4F). These results support the idea that

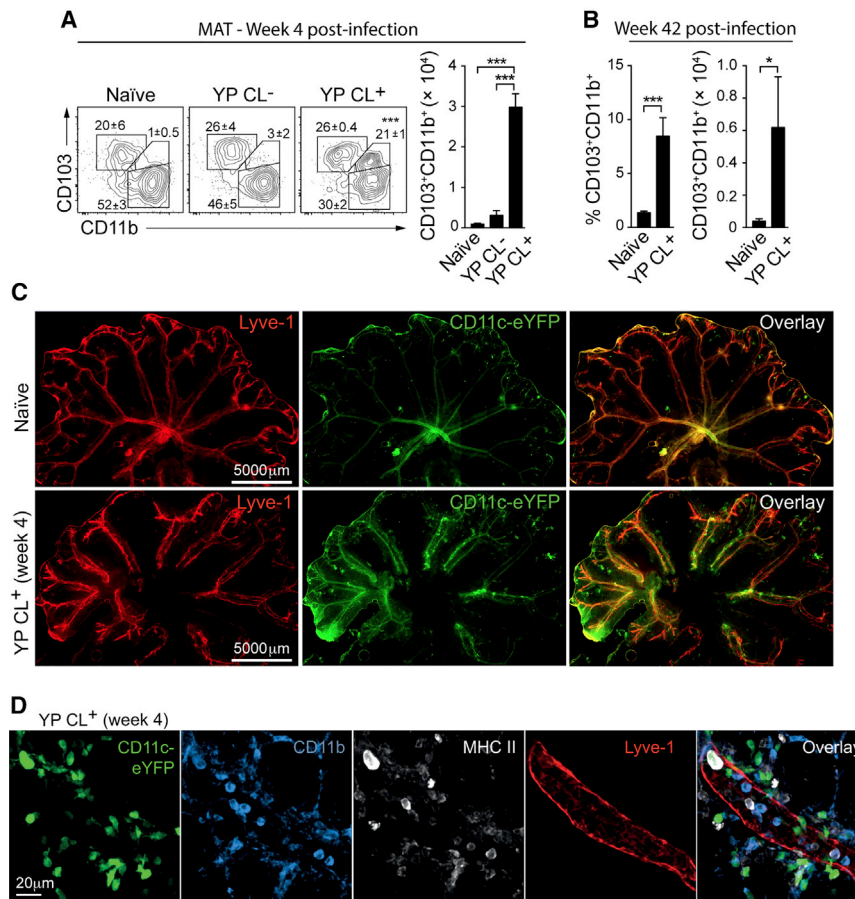


Figure 5. CD103⁺CD11b⁺ DCs Accumulate in the MAT following Infection with *Yersinia pseudotuberculosis*

(A) DC subsets in the MAT as characterized by flow cytometry, according to the gate strategy described in Figure S4E. Shown are representative contour plots of MAT DC subsets from naïve or infected animals (CL^{+/−}). Numbers in plot indicate the mean frequency of cells (±SEM) within the adjacent gate from all samples in this experiment. Bar graph shows the number of CD103⁺CD11b⁺ DCs in the MAT.

(B) Bar graphs show the frequency and number of CD103⁺CD11b⁺ DCs in the MAT at 42 weeks post-infection.

(C) Representative whole tissue fluorescent images of MAT from naïve or 4-week infected CD11c-YFP reporter mice.

(D) Representative confocal images describing the localization of CD11c⁺CD11b⁺MHC II⁺ cells in relation to lymphatics (LYVE1⁺) in the MAT. All the data are representative of three experiments each containing five control animals and five to ten infected animals. All bar graphs show the mean (±SEM). *p < 0.05, ***p < 0.0005 (Student's t test). CL⁺, infected mice with lymphadenopathy; CL[−], infected mice without lymphadenopathy. See also Figure S4.

following infection, migratory DCs prematurely escape the lymphatic flow and are shunted to the adipose tissue.

Microbiota Sustains Impaired Tissue-Specific Immunity following Infection

We next explored the possibility that the resident microbiota might sustain tissue remodeling and inflammation following resolution of the infection. To address this possibility, we assessed the consequences of *Y. pseudotuberculosis* infection in mice devoid of live microbes (germ-free [GF]). Like specific pathogen-free (SPF) mice, GF mice cleared the bacteria from the MLN, liver, and spleen, but, in contrast to SPF mice, not in the GI tract (Figures S5A and S5B). This observation is consistent with the known role for the microbiota as an agent of colonization resistance (Buffie and Pamer, 2013). GF mice still developed lymphadenopathy, albeit to a lesser magnitude than in SPF mice, revealing that the microbiota was not strictly required for infection-induced remodeling of the MLN (Figure 6A). On the other hand, *Y. pseudotuberculosis* had no impact on the capacity of CD103⁺CD11b⁺ DCs to accumulate in the MLN of GF mice, highlighting a central role for the microbiota in the induction and/or maintenance of this specific sequela of infection (Figure 6B).

We next evaluated whether a transient ablation of the microbiota in SPF mice could reduce MAT inflammation and restore

mucosal immune function. Antibiotic treatment significantly reduced type I inflammation of the MAT, prevented the aberrant accumulation of CD103⁺CD11b⁺ DCs in the MAT, and restored their physiological homing to the MLN (Figures 6C and 6D). However, while MAT inflammation was reduced following antibiotic treatment, lymphatic integrity was not fully restored (Figure S5C). Oral gavage of such mice with microbial-derived products was sufficient to restore cellular infiltration into the MAT, (Figures S5D and S5E) suggesting that the effect may be mediated by enhanced microbial exposure independent of any defined microbiota. Indeed, mice developing lymphadenopathy can be found in the same cage as CL[−] mice (Figure S5F).

We next assessed if antibiotic treatment was sufficient to restore mucosal immunity. As previously described, antibiotic treatment reduced the efficacy of oral vaccination (Hall et al., 2008) (Figures 6E and S5G). Nonetheless, a substantial number of Th17 cells specific for dmlT or OVA were still induced and accumulated in the lamina propria of vaccinated mice (Figure 6E). Together with restoration of MAT homeostasis, antibiotic treatment also restored the capacity of CL⁺ mice to respond to oral vaccination to a level comparable to naïve antibiotic-treated mice (Figure 6E). Thus, damage induced by acute infections may be reversible and limiting exposure to defined microbial products while promoting lymphatic repair may represent a therapeutic strategy to restore mucosal immunity. Taken together, these data reveal a central role for the microbiota in the maintenance of the immune defects induced by acute infection.

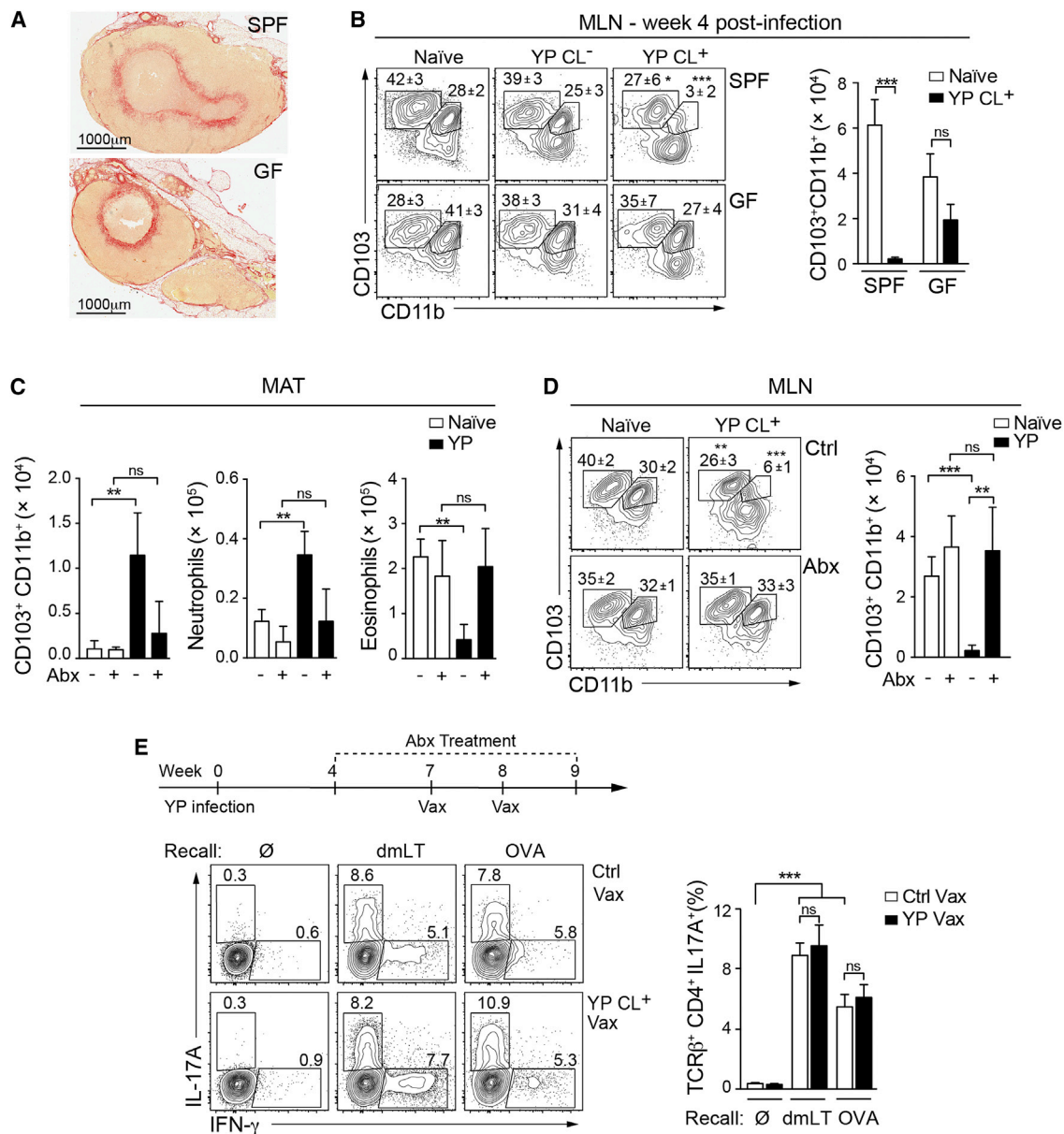


Figure 6. Microbiota Sustains MAT Inflammation and Immune Dysfunction Post-infection

(A and B) Specific pathogen-free (SPF) or germ-free (GF) C57BL/6 mice were infected with *Y. pseudotuberculosis*. (A) Sections of MLNs isolated 4 weeks post-infection were stained with picosirius red. (B) Frequency and number of DC subsets in the MLN. Shown are representative flow cytometric contour plots of DC subsets (gated according to the Figure S2C). Numbers in plots represent the mean frequency (\pm SEM) of cells within the adjacent gate. Bar graph shows the numbers of CD103⁺CD11b⁺ DCs in the MLN.

(C and D) Four weeks post-infection, SPF mice were treated with broad-spectrum antibiotics (Abx) for 3 weeks. (C) Bar graphs show the number of CD103⁺CD11b⁺ DCs (gated as indicated in Figure S4E), neutrophils, and eosinophils isolated from the MAT of naïve or infected animals with or without antibiotic treatment. (D) Shown are representative plots of the MLN DC subsets from naïve and infected mice, with and without antibiotic treatment. Numbers represent the mean frequency (\pm SEM) of cells within the adjacent gate of all samples within this experiment. Bar graph shows the number of CD103⁺CD11b⁺ DCs in the MLNs.

(E) Top: scheme for vaccination of antibiotic-treated mice. Abx treatment was continued during the vaccination period. 7 days post-immunization, lymphocytes were isolated from the siLP and stimulated in vitro with dmLT or OVA pulsed DCs. Bottom left: contour plots showing IFN- γ and IL-17A producing CD4⁺ T cells after vaccination in naïve (Ctrl Vax) or infected (YP Vax) mice treated with Abx. Numbers represent the percent of CD4⁺ T cells that express each cytokine in the adjacent gate. The bar graph (bottom right) shows the mean percent of IL-17A producing CD4⁺ T cells from Abx treated naïve or infected mice. All data are representative of two to three experiments with three to five mice in each control group and five to seven mice in each infected group. All bar graphs show the mean (\pm SEM). * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$ (Student's *t* test). CL⁺, infected mice with lymphadenopathy; CL⁻, infected mice without lymphadenopathy; ns, not significant.

See also Figure S5.

DISCUSSION

Here, we report that a single acute infection can lead to long-lived “immunological scarring” with profound effects on tissue immunity that persist after the pathogenic organism has been cleared. Infections have long been considered as possible triggering elements in the development of chronic inflammatory diseases and immune dysfunction, but definitive and mechanistic associations between defined pathogens and the breakdown of tissue homeostasis have remained elusive. Our work provides a mechanistic framework via which infection may represent a point of no return for tissue immunity, leading to chronic inflammation and immune dysfunction long after the pathogen has been cleared.

In this study, we found that oral infection with *Y. pseudotuberculosis* profoundly and persistently remodels both the MAT and MLN. The pathological effect of infection in this mouse model reproduces the chronic lymphadenopathy that is a defining characteristic of *Y. enterocolitica* or *Y. pseudotuberculosis* infection in humans (Asano, 2012). Chronic mesenteric lymphadenopathy is not limited to *Yersinia* infection, as a variety of other diseases can also cause these symptoms, such as HIV/AIDS, Crohn’s disease, and celiac disease (Asano, 2012; Estes et al., 2008; Huppert et al., 2004; Lucey et al., 2005). The mechanism by which *Y. pseudotuberculosis* triggers damage to the lymphatic system is not understood, but is possibly related to the fact that *Yersinia* and other pathogens drive phagocytic cells to travel through these vessels (Bou Ghanem et al., 2012; Diehl et al., 2013; St John et al., 2014). We also show that increased permeability of lymphatics can be observed in other models of GI infection or non-infectious colitis. As such, the point raised by our present work is not that *Y. pseudotuberculosis* per se is responsible for the breakdown of tissue immunity, but that in defined settings or in genetically predisposed individuals, GI infections and/or inflammation, may play a determinant role in the disruption of immune homeostasis by remodeling the immune and lymphatic system.

Two cytokines, IL-1 β and TNF- α , have been shown to increase lymphatic permeability and both are significantly increased in the MAT post-infection (Aldrich and Sevick-Muraca, 2013; Cromer et al., 2014). Another potential contributing factor in chronic inflammation of the MAT is the switch from type 2 to type 1 immunity. Because type 2 immunity has recently been described as critical to tissue healing (Gause et al., 2013), persistent lymphatic damage post-infection may also be the consequence of a loss of these responses. There are a number of interesting parallels between the chronic remodeling of the MAT seen after *Y. pseudotuberculosis* infection and inflammatory bowel disease. Notably, inflammation of the MAT, altered lymph drainage and presence of translocating commensals in this compartment has been described in Crohn’s disease patients (Behr, 2010; Peyrin-Biroulet et al., 2012; Zulian et al., 2013). In accordance with these studies, we measured increased lymphatic leakage in the T cell transfer model of colitis. Previous studies have also shown that the larger, contractile lymphatics are permeable to soluble antigens and that adipose tissue dendritic cells can sample molecules traveling in the lymph (Kuan et al., 2015). Furthermore, when permeability of lymphatic collecting vessels

is increased, the absorptive function of upstream lymphatic capillaries, such as those in intestinal villi, is decreased (Scallan et al., 2013) a phenomenon that could be relevant to our present observation. Therefore, exploration of the involvement of lymphatic damage and inflammation in the MAT could be important to our understanding of the pathophysiology of Crohn’s disease and gut inflammatory disorders in general.

Infection abrogated the accumulation of the migratory CD103⁺CD11b⁺ DCs in the MLNs, a cardinal subset of gut DCs involved in the promotion of both T and B cell responses, (Bogunovic et al., 2009; Farache et al., 2013; Satpathy et al., 2013; Schlitzer et al., 2013). We also observed a partial reduction in the frequency of CD103⁺CD11b⁺ DCs in the MLN, a population associated with T_{reg} induction (Coombes et al., 2007; Sun et al., 2007). Because inflammation can complicate the identification of these cells (Merad et al., 2013), migratory CD103⁺CD11b⁺ DCs may be more profoundly affected by the infection than reported here, a point supported by the substantial impact of the infection on T_{reg} induction. Of interest, infection-induced disruption of the positioning of DCs and macrophages in the LN has been observed in multiple settings of infection and vaccination and similar defects due to lymphadenopathy may further affect DC function in our model (Gaya et al., 2015; Katakai et al., 2004; Mueller et al., 2007).

Immunological damage from acute infection had profound consequences for the maintenance of tissue-specific immunity, including defects in the development of Th17 cells, T_{reg}, and IgA⁺ B cell responses in the GI tract. Collectively, these responses represent the canonical mucosal immune response required for oral tolerance, adaptive immunity, and compartmentalization of the microbiota (Belkaid and Hand, 2014). Our findings indicate that under conditions in which the environment is in flux, some acute infections may predispose to food allergy and aberrant immune responses to commensals via the abrogation of T_{reg} induction. In low to middle income countries, oral vaccines for Rotavirus, Poliomyelitis, *Vibrio cholerae*, and *Shigella* often fail to protect with the same degree of efficacy as in high-income countries (Levine, 2010). Several interrelated conditions, including persistent infections, gut inflammation, malnutrition, and environmental enteropathy, have been proposed to contribute to this important public health issue (Korpe and Petri, 2012; Levine, 2010). Our observed defects in DC trafficking and the development of Th17 and IgA responses to model oral vaccinations may provide an explanation for the immune dysfunction in these populations. Interestingly, when individuals suffering from environmental enteropathy are relocated to Western countries, where diet and sanitation are no longer issues, it can take up to 2 years for intestinal absorption to be restored, implying that the physiological defects outlive exposure to acute infections (Lindenbaum et al., 1972). Based on our present work, we propose that the integrity of the MAT and lymphatics could be also severely affected under these settings and that this phenomenon may contribute to immune dysfunction.

The microbiota plays a central role in the maintenance of the immunological damage caused by *Y. pseudotuberculosis* infection. This observation reveals yet another way by which interruption of the host’s relationship with its microbiota can compromise tissue homeostasis. Our current model proposes that

such effects may not be associated with a defined microbiota. Notably, mice developing lymphadenopathy do not have a unique microbiota compared to those mice that did not develop lymphadenopathy and oral gavage of defined microbial products can phenocopy the effect of the microbiota. Thus, increased exposure to microbes or microbe-derived ligands provided by the increased leakage of the lymphatics may be sufficient to induce a positive feedback loop of inflammation in the MAT, an effect that could predispose to the development of inflammatory disorders. On the other hand, we cannot exclude the possibility that benign microbes such as lactobacilli that can be found in the CL⁺ MLN could be contextually pathogenic post-infection.

The “hygiene hypothesis” proposes that increasing rates of allergy and asthma in western countries could be the consequence of reduced infectious disease, in particular helminthic worm infection during early childhood (Johnston et al., 2014; Schaub et al., 2006; Weinstock et al., 2004). More recently, this model has been modified to include the modern shift in the human microbiota (Blaser and Falkow, 2009). Together, these profound changes in microbial encounters and, as a direct result, the state of the immune system, are believed to contribute to the dramatic increase in chronic inflammatory and autoimmune disorders seen in high-income countries. Our present work proposes that this model also needs to integrate the change in the frequency and type of acute infections associated with high-density urban living. Modern societies are now burdened with endemic respiratory and GI infections, whose long-term sequelae are not fully understood. These repeated and unregulated inflammatory challenges may profoundly remodel the immune system and thereby contribute to the increased burden of autoimmune and inflammatory disorders.

Together, this study provides a framework to understand how previously encountered infections can induce a breakdown of tissue immune homeostasis, thereby contributing to disease later in life. Thus, in order to fully comprehend the etiology of complex diseases, it may be necessary to look beyond a patient's genetic susceptibilities and concurrent environmental stressors and examine whether immunological scarring associated with previous infections may have “set the stage” for chronic inflammation.

EXPERIMENTAL PROCEDURES

Mice

C57BL/6NTac SPF mice were purchased from Taconic Farms. Germ-free C57BL/6 mice were bred at Taconic Farms and maintained in the National Institute of Allergy and Infectious Diseases (NIAID) gnotobiotic facility. C57BL/6NTac-Lysozyme (Lyz2) (LysM-eGFP reporter), B6.SJL-Ptprc^a/BoyAiTac (CD45.1), C57BL/6-[Tg]CD11c(*Itgax*):EYFP (CD11c-eYFP reporter), B6.129P(Cg)-Ptprc^a Cx3cr1^{tm1Litt}/LittJ (CX₃CR1 GFP reporter), and C57BL/6-CD45a(Ly5a)-*Rag1*^{-/-} TCR OT-II (*Rag1*^{-/-} OT-II TCR transgenic) mice were obtained through the NIAID-Taconic exchange program. All mice were bred and maintained under pathogen-free conditions at an American Association for the Accreditation of Laboratory Animal Care accredited animal facility at the NIAID and housed in accordance with the procedures outlined in the Guide for the Care and Use of Laboratory Animals. Gender- and age-matched mice between 6–12 weeks old were used. When possible, preliminary experiments were performed to determine requirements for sample size, taking into account resources available and ethical, reductionist animal use. In general, each mouse of the different experimental groups is reported. Exclusion criteria

such as inadequate staining or low cell yield due to technical problems were pre-determined.

Oral Infection, Vaccination, and Treatments

For infection, *Y. pseudotuberculosis* (strain IP32777) (Simonet and Falkow, 1992) was grown into 2XYT media (Quality Biological) overnight at 25°C with vigorous shaking. Mice were fasted for 12 hr prior to infection with 1×10^7 CFU via oral gavage. Oral vaccination with double mutant *E. coli* heat labile toxin (R192G/L211A) (dmLT) (provided by J. Clements) was carried out as previously described (Hall et al., 2008). Mice received two immunizations (7 days apart and response was assessed 7 days after). For broad-spectrum antibiotic treatment, mice were treated as previously described (Hall et al., 2008). Commensal-derived DNA/LPS gavage of antibiotic-treated mice was performed as previously described (Hall et al., 2008), with the modification that 250 µg of commensal DNA and 250 µg of LPS (from *E. coli*; Sigma) were gavaged every other day for a week.

Analysis of Leaking in the MAT

Bodipy FL C16 (40 µg/mouse) (Life Technologies) mixed with milk cream (10%) (100 µl) and olive oil (100 µl) was gavaged into mice and MAT tissue was isolated for microscopy 3 hr later and analyzed by scanning fluorescent microscopy (Leica M205 FA, motorized stereo microscope) and confocal microscopy (Leica SP8).

Statistics

Data are presented as mean ± SEM. Group sizes were determined based on the results of preliminary experiments. Mice were assigned at random to groups. Mouse studies were not performed in a blinded fashion. Statistical significance was determined with the two-tailed unpaired Student's *t* test, under the untested assumption of normality. Within each group there was an estimate of variation, and the variance between groups was similar. All statistical analysis was calculated using Prism software (GraphPad). Differences were considered to be statistically significant when *p* < 0.05.

For more detailed information, please consult the [Supplemental Experimental Procedures](#).

ACCESSION NUMBERS

The accession number for the 16S sequencing data reported in this paper is SRA: PRJNA276477.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.030>.

AUTHOR CONTRIBUTIONS

D.M.F., T.W.H., and Y.B. designed the studies and wrote the manuscript. D.M.F. and T.W.H. performed the experiments and analyzed the data. D.M.F., S.J.H., and M.Y.G. generated and analyzed confocal imaging data. O.J.H. and A.G.-Z. participated by performing experiments. G.T. assisted with designing the NanoString gene expression. J.M.B. and A.M.O. assisted with experiments on non-human primates. A.L.B. performed NanoString data analysis. M.Q. analyzed 16S sequencing data. I.E.B. provided the *Y. pseudotuberculosis* and contributed with guidance to establish the model. R.N.G. and G.J.R. provided intellectual expertise and helped to interpret experimental results.

ACKNOWLEDGMENTS

This work was supported by the Division of Intramural Research of the National Institute of Allergy and Infectious Diseases (NIAID), NIH. D.M.F. was supported by FAPESP (2012/14669-7) and CNPq/CAPES (10986-13-8). T.W.H. was supported by a FDA/ODS Scholarship. We thank J. D. Clements (University of

Tulane School of Medicine) for providing the *E. coli* dMLT. We thank D. Trageser-Cesler and C. Acevedo (NIAID Gnotobiotic Animal Facility); NIAID animal facility staff; K. Holmes, C. Eigsti, T. Moyer, and E. Stregevsy (NIAID Flow Cytometry facility); O. Schwartz, J. Kabat, L. Koo, and M. Smelkinson (NIAID Biological Imaging facility) and N. Bubunenko, D. Sun, R. Winkler-Pickett, K. Beacht, J. Davis, M. Solano-Dias and J. Legrand for technical assistance. We thank N. Bouladoux, S. Henri, A. Poholek, M. Constantinides, V. Ridaura and all the members of the Y.B. laboratory for critical reading of the manuscript and helpful discussions.

Received: February 26, 2015

Revised: July 9, 2015

Accepted: August 11, 2015

Published: October 8, 2015

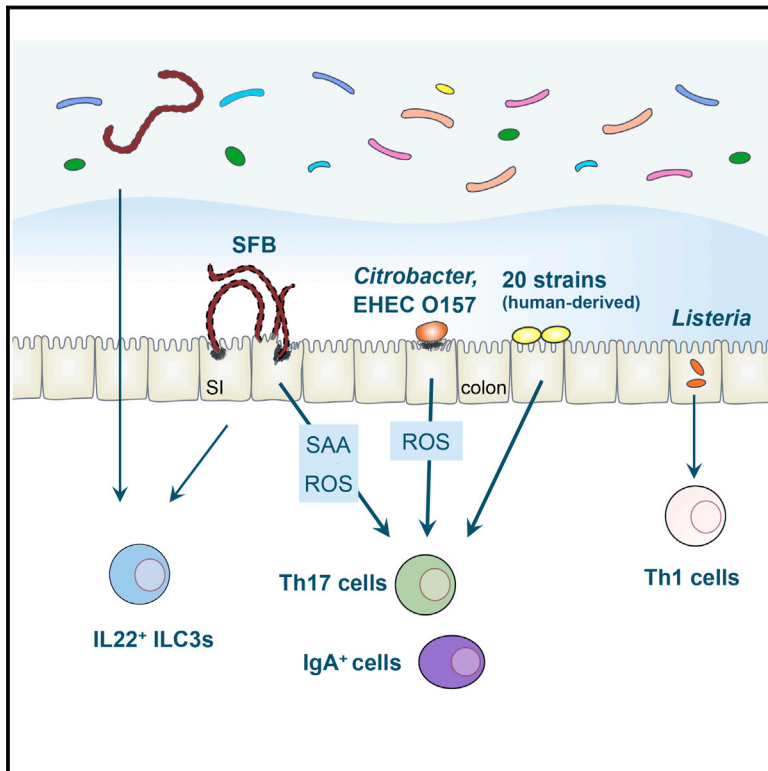
REFERENCES

- Aldrich, M.B., and Sevrick-Muraca, E.M. (2013). Cytokines are systemic effectors of lymphatic function in acute inflammation. *Cytokine* 64, 362–369.
- Asano, S. (2012). Granulomatous lymphadenitis. *J. Clin. Exp. Hematop.* 52, 1–16.
- Barton, E.S., White, D.W., Cathelyn, J.S., Brett-McClellan, K.A., Engle, M., Diamond, M.S., Miller, V.L., and Virgin, H.W., 4th. (2007). Herpesvirus latency confers symbiotic protection from bacterial infection. *Nature* 447, 326–329.
- Behr, M.A. (2010). The path to Crohn's disease: is mucosal pathology a secondary event? *Inflamm. Bowel Dis.* 16, 896–902.
- Bekiaris, V., Persson, E.K., and Agace, W.W. (2014). Intestinal dendritic cells in the regulation of mucosal immunity. *Immunol. Rev.* 260, 86–101.
- Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. *Cell* 157, 121–141.
- Belkaid, Y., and Segre, J.A. (2014). Dialogue between skin microbiota and immunity. *Science* 346, 954–959.
- Bergman, M.A., Loomis, W.P., Mecsas, J., Starnbach, M.N., and Isberg, R.R. (2009). CD8(+) T cells restrict *Yersinia pseudotuberculosis* infection: bypass of anti-phagocytosis by targeting antigen-presenting cells. *PLoS Pathog.* 5, e1000573.
- Bergsbaken, T., and Bevan, M.J. (2015). Proinflammatory microenvironments within the intestine regulate the differentiation of tissue-resident CD8 T cells responding to infection. *Nat. Immunol.*
- Blaser, M.J., and Falkow, S. (2009). What are the consequences of the disappearing human microbiota? *Nat. Rev. Microbiol.* 7, 887–894.
- Bogunovic, M., Ginhoux, F., Helft, J., Shang, L., Hashimoto, D., Greter, M., Liu, K., Jakubczak, C., Ingersoll, M.A., Leboeuf, M., et al. (2009). Origin of the lamina propria dendritic cell network. *Immunity* 31, 513–525.
- Bou Ghanem, E.N., Jones, G.S., Myers-Morales, T., Patil, P.D., Hidayatullah, A.N., and D'Orazio, S.E. (2012). InlA promotes dissemination of *Listeria monocytogenes* to the mesenteric lymph nodes during food borne infection of mice. *PLoS Pathog.* 8, e1003015.
- Buffie, C.G., and Pamer, E.G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* 13, 790–801.
- Cadwell, K., Patel, K.K., Maloney, N.S., Liu, T.C., Ng, A.C., Storer, C.E., Head, R.D., Xavier, R., Stappenbeck, T.S., and Virgin, H.W. (2010). Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell* 141, 1135–1145.
- Coombes, J.L., Siddiqui, K.R., Arancibia-Carcamo, C.V., Hall, J., Sun, C.M., Belkaid, Y., and Powrie, F. (2007). A functionally specialized population of mucosal CD103+ DCs induces Foxp3+ regulatory T cells via a TGF- β and retinoic acid-dependent mechanism. *J. Exp. Med.*
- Cromer, W.E., Zawieja, S.D., Tharakan, B., Childs, E.W., Newell, M.K., and Zawieja, D.C. (2014). The effects of inflammatory cytokines on lymphatic endothelial barrier function. *Angiogenesis* 17, 395–406.
- Curotto de Lafaille, M.A., Kutchukhidze, N., Shen, S., Ding, Y., Yee, H., and Lafaille, J.J. (2008). Adaptive Foxp3+ regulatory T cell-dependent and -independent control of allergic inflammation. *Immunity* 29, 114–126.
- DePaolo, R.W., Abadie, V., Tang, F., Fehlner-Peach, H., Hall, J.A., Wang, W., Marietta, E.V., Kasarda, D.D., Waldmann, T.A., Murray, J.A., et al. (2011). Co-adjuvant effects of retinoic acid and IL-15 induce inflammatory immunity to dietary antigens. *Nature* 471, 220–224.
- Diehl, G.E., Longman, R.S., Zhang, J.X., Breart, B., Galan, C., Cuesta, A., Schwab, S.R., and Littman, D.R. (2013). Microbiota restricts trafficking of bacteria to mesenteric lymph nodes by CX(3)CR1(hi) cells. *Nature* 494, 116–120.
- Estes, J., Baker, J.V., Brenchley, J.M., Khoruts, A., Barthold, J.L., Bantle, A., Reilly, C.S., Beilman, G.J., George, M.E., Douek, D.C., et al. (2008). Collagen deposition limits immune reconstitution in the gut. *J. Infect. Dis.* 198, 456–464.
- Fahlgren, A., Avican, K., Westermarck, L., Nordfelth, R., and Fällman, M. (2014). Colonization of cecum is important for development of persistent infection by *Yersinia pseudotuberculosis*. *Infect. Immun.* 82, 3471–3482.
- Farache, J., Koren, I., Milo, I., Gurevich, I., Kim, K.W., Zigmond, E., Furtado, G.C., Lira, S.A., and Shakhar, G. (2013). Luminal bacteria recruit CD103+ dendritic cells into the intestinal epithelium to sample bacterial antigens for presentation. *Immunity* 38, 581–595.
- Fendrick, A.M., Monto, A.S., Nightengale, B., and Sarnes, M. (2003). The economic burden of non-influenza-related viral respiratory tract infection in the United States. *Arch. Intern. Med.* 163, 487–494.
- Gause, W.C., Wynn, T.A., and Allen, J.E. (2013). Type 2 immunity and wound healing: evolutionary refinement of adaptive immunity by helminths. *Nat. Rev. Immunol.* 13, 607–614.
- Gaya, M., Castello, A., Montaner, B., Rogers, N., Reis e Sousa, C., Bruckbauer, A., and Batista, F.D. (2015). Host response. Inflammation-induced disruption of SCS macrophages impairs B cell responses to secondary infection. *Science* 347, 667–672.
- Gerner, M.Y., Kastentmuller, W., Ifrim, I., Kabat, J., and Germain, R.N. (2012). Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* 37, 364–376.
- Grainger, J.R., Askenase, M.H., Guimont-Desrochers, F., da Fonseca, D.M., and Belkaid, Y. (2014). Contextual functions of antigen-presenting cells in the gastrointestinal tract. *Immunol. Rev.* 259, 75–87.
- Hadis, U., Wahl, B., Schulz, O., Hardtke-Wolenski, M., Schippers, A., Wagner, N., Müller, W., Sparwasser, T., Förster, R., and Pabst, O. (2011). Intestinal tolerance requires gut homing and expansion of FoxP3+ regulatory T cells in the lamina propria. *Immunity* 34, 237–246.
- Hall, J.A., Bouladoux, N., Sun, C.M., Wohlfert, E.A., Blank, R.B., Zhu, Q., Grigg, M.E., Berzofsky, J.A., and Belkaid, Y. (2008). Commensal DNA limits regulatory T cell conversion and is a natural adjuvant of intestinal immune responses. *Immunity* 29, 637–649.
- Hand, T.W., Dos Santos, L.M., Bouladoux, N., Molloy, M.J., Pagán, A.J., Pepper, M., Maynard, C.L., Elson, C.O., 3rd, and Belkaid, Y. (2012). Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. *Science* 337, 1553–1556.
- Hooper, L.V., and Macpherson, A.J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* 10, 159–169.
- Huppert, B.J., Farrell, M.A., Kawashima, A., and Murray, J.A. (2004). Diagnosis of cavitating mesenteric lymph node syndrome in celiac disease using MRI. *AJR Am. J. Roentgenol.* 183, 1375–1377.
- Jamieson, A.M., Pasman, L., Yu, S., Gamradt, P., Homer, R.J., Decker, T., and Medzhitov, R. (2013). Role of tissue protection in lethal respiratory viral-bacterial coinfection. *Science* 340, 1230–1234.
- Johnston, C.J., McSorley, H.J., Anderton, S.M., Wigmore, S.J., and Maizels, R.M. (2014). Helminths and immunological tolerance. *Transplantation* 97, 127–132.
- Katakai, T., Hara, T., Sugai, M., Gonda, H., and Shimizu, A. (2004). Lymph node fibroblastic reticular cells construct the stromal reticulum via contact with lymphocytes. *J. Exp. Med.* 200, 783–795.
- Khalifeh-Soltani, A., McKleroy, W., Sakuma, S., Cheung, Y.Y., Tharp, K., Qiu, Y., Turner, S.M., Chawla, A., Stahl, A., and Atabai, K. (2014). Mfge8 promotes

- obesity by mediating the uptake of dietary fats and serum fatty acids. *Nat. Med.* 20, 175–183.
- Kissenpfennig, A., Henri, S., Dubois, B., Laplace-Builhé, C., Perrin, P., Romani, N., Tripp, C.H., Douillard, P., Leserman, L., Kaiserlian, D., et al. (2005). Dynamics and function of Langerhans cells in vivo: dermal dendritic cells colonize lymph node areas distinct from slower migrating Langerhans cells. *Immunity* 22, 643–654.
- Korpe, P.S., and Petri, W.A., Jr. (2012). Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol. Med.* 18, 328–336.
- Kosek, M., Bern, C., and Guerrant, R.L. (2003). The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull. World Health Organ.* 81, 197–204.
- Kuan, E.L., Ivanov, S., Bridenbaugh, E.A., Victora, G., Wang, W., Childs, E.W., Platt, A.M., Jakubzick, C.V., Mason, R.J., Gashev, A.A., et al. (2015). Collecting lymphatic vessel permeability facilitates adipose tissue inflammation and distribution of antigen to lymph node-homing adipose tissue dendritic cells. *J. Immunol.* 194, 5200–5210.
- Levine, M.M. (2010). Immunogenicity and efficacy of oral vaccines in developing countries: lessons from a live cholera vaccine. *BMC Biol.* 8, 129.
- Lindenbaum, J., Harmon, J.W., and Gerson, C.D. (1972). Subclinical malabsorption in developing countries. *Am. J. Clin. Nutr.* 25, 1056–1061.
- Lucey, B.C., Stuhlfaut, J.W., and Soto, J.A. (2005). Mesenteric lymph nodes seen at imaging: causes and significance. *Radiographics* 25, 351–365.
- Maloy, K.J., and Powrie, F. (2011). Intestinal homeostasis and its breakdown in inflammatory bowel disease. *Nature* 474, 298–306.
- McPhee, J.B., Mena, P., Zhang, Y., and Bliska, J.B. (2012). Interleukin-10 induction is an important virulence function of the *Yersinia pseudotuberculosis* type III effector YopM. *Infect. Immun.* 80, 2519–2527.
- Merad, M., Sathe, P., Helft, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu. Rev. Immunol.* 31, 563–604.
- Molodecky, N.A., Soon, I.S., Rabi, D.M., Ghali, W.A., Ferris, M., Chernoff, G., Benchimol, E.I., Panaccione, R., Ghosh, S., Barkema, H.W., et al. (2012). Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 142, 46–54, e42; quiz e30.
- Mowat, A.M., and Agace, W.W. (2014). Regional specialization within the intestinal immune system. *Nat. Rev. Immunol.* 14, 667–685.
- Mueller, S.N., Matloubian, M., Clemens, D.M., Sharpe, A.H., Freeman, G.J., Gangappa, S., Larsen, C.P., and Ahmed, R. (2007). Viral targeting of fibroblastic reticular cells contributes to immunosuppression and persistence during chronic infection. *Proc. Natl. Acad. Sci. USA* 104, 15430–15435.
- Oldenhove, G., Bouladoux, N., Wohlfert, E.A., Hall, J.A., Chou, D., Dos Santos, L., O'Brien, S., Blank, R., Lamb, E., Natarajan, S., et al. (2009). Decrease of Foxp3(+) Treg Cell Number and Acquisition of Effector Cell Phenotype during Lethal Infection. *Immunity*.
- Osborne, L.C., Monticelli, L.A., Nice, T.J., Sutherland, T.E., Siracusa, M.C., Hepworth, M.R., Tomov, V.T., Kobuley, D., Tran, S.V., Bittinger, K., et al. (2014). Coinfection. Virus-helminth coinfection reveals a microbiota-independent mechanism of immunomodulation. *Science* 345, 578–582.
- Peyrin-Biroulet, L., Gonzalez, F., Dubuquoy, L., Rousseaux, C., Dubuquoy, C., Decourcelle, C., Saudemont, A., Tachon, M., Béclin, E., Odou, M.F., et al. (2012). Mesenteric fat as a source of C reactive protein and as a target for bacterial translocation in Crohn's disease. *Gut* 61, 78–85.
- Randolph, G.J., and Miller, N.E. (2014). Lymphatic transport of high-density lipoproteins and chylomicrons. *J. Clin. Invest.* 124, 929–935.
- Salgame, P., Yap, G.S., and Gause, W.C. (2013). Effect of helminth-induced immunity on infections with microbial pathogens. *Nat. Immunol.* 14, 1118–1126.
- Satpathy, A.T., Briseño, C.G., Lee, J.S., Ng, D., Manieri, N.A., Kc, W., Wu, X., Thomas, S.R., Lee, W.L., Turkoz, M., et al. (2013). Notch2-dependent classical dendritic cells orchestrate intestinal immunity to attaching-and-effacing bacterial pathogens. *Nat. Immunol.* 14, 937–948.
- Scallan, J.P., Davis, M.J., and Huxley, V.H. (2013). Permeability and contractile responses of collecting lymphatic vessels elicited by atrial and brain natriuretic peptides. *J. Physiol.* 591, 5071–5081.
- Schaub, B., Lauener, R., and von Mutius, E. (2006). The many faces of the hygiene hypothesis. *J. Allergy Clin. Immunol.* 117, 969–977, quiz 978.
- Schlitzer, A., McGovern, N., Teo, P., Zelante, T., Atarashi, K., Low, D., Ho, A.W., See, P., Shin, A., Wasan, P.S., et al. (2013). IRF4 transcription factor-dependent CD11b+ dendritic cells in human and mouse control mucosal IL-17 cytokine responses. *Immunity* 38, 970–983.
- Schulz, O., Jaensson, E., Persson, E.K., Liu, X., Worbs, T., Agace, W.W., and Pabst, O. (2009). Intestinal CD103+, but not CX3CR1+, antigen sampling cells migrate in lymph and serve classical dendritic cell functions. *J. Exp. Med.* 206, 3101–3114.
- Simonet, M., and Falkow, S. (1992). Invasin expression in *Yersinia pseudotuberculosis*. *Infect. Immun.* 60, 4414–4417.
- St John, A.L., Ang, W.X., Huang, M.N., Kunder, C.A., Chan, E.W., Gunn, M.D., and Abraham, S.N. (2014). S1P-Dependent trafficking of intracellular yersinia pestis through lymph nodes establishes Buboes and systemic infection. *Immunity* 41, 440–450.
- Sun, C.M., Hall, J.A., Blank, R.B., Bouladoux, N., Oukka, M., Mora, J.R., and Belkaid, Y. (2007). Small intestine lamina propria dendritic cells promote de novo generation of Foxp3 T reg cells via retinoic acid. *J. Exp. Med.* 204, 1775–1785.
- Varol, C., Vallon-Eberhard, A., Elinav, E., Aycheh, T., Shapira, Y., Lucche, H., Fehling, H.J., Hardt, W.D., Shakhar, G., and Jung, S. (2009). Intestinal lamina propria dendritic cell subsets have different origin and functions. *Immunity* 31, 502–512.
- Vernacchio, L., Vezina, R.M., Mitchell, A.A., Lesko, S.M., Plaut, A.G., and Acheson, D.W. (2006). Diarrhea in American infants and young children in the community setting: incidence, clinical presentation and microbiology. *Pediatr. Infect. Dis. J.* 25, 2–7.
- Wang, Y., Su, M.A., and Wan, Y.Y. (2011). An essential role of the transcription factor GATA-3 for the function of regulatory T cells. *Immunity* 35, 337–348.
- Weiner, H.L., da Cunha, A.P., Quintana, F., and Wu, H. (2011). Oral tolerance. *Immunol. Rev.* 241, 241–259.
- Weinstock, J.V., Summers, R., and Elliott, D.E. (2004). Helminths and harmony. *Gut* 53, 7–9.
- Wohlfert, E.A., Grainger, J.R., Bouladoux, N., Konkel, J.E., Oldenhove, G., Ribeiro, C.H., Hall, J.A., Yagi, R., Naik, S., Bhairavabhotla, R., et al. (2011). GATA3 controls Foxp3+ regulatory T cell fate during inflammation in mice. *J. Clin. Invest.* 121, 4503–4515.
- Worbs, T., Bode, U., Yan, S., Hoffmann, M.W., Hintzen, G., Bernhardt, G., Förster, R., and Pabst, O. (2006). Oral tolerance originates in the intestinal immune system and relies on antigen carriage by dendritic cells. *J. Exp. Med.* 203, 519–527.
- Wren, B.W. (2003). The yersiniae—a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* 1, 55–64.
- Zhang, Y., Mena, P., Romanov, G., Lin, J.S., Smiley, S.T., and Bliska, J.B. (2012). A protective epitope in type III effector YopE is a major CD8 T cell antigen during primary infection with *Yersinia pseudotuberculosis*. *Infect. Immun.* 80, 206–214.
- Zulian, A., Cancellato, R., Ruocco, C., Gentilini, D., Di Blasio, A.M., Danelli, P., Micheletto, G., Cesana, E., and Invitti, C. (2013). Differences in visceral fat and fat bacterial colonization between ulcerative colitis and Crohn's disease. An in vivo and in vitro study. *PLoS ONE* 8, e78495.

Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells

Graphical Abstract



Authors

Koji Atarashi, Takeshi Tanoue, Minoru Ando, ..., Masahira Hattori, Yoshinori Umesaki, Kenya Honda

Correspondence

yoshinori-umesaki@yakult.co.jp (Y.U.), kenya@keio.jp (K.H.)

In Brief

The adhesion of specific members of the gut microbiome to intestinal epithelial cells is found to be essential for the induction of Th17 cells, thus highlighting location as an additional layer of regulation for the adaptive immune system.

Highlights

- A strong correlation between epithelial adhesion and Th17 induction by SFB and EHEC
- Twenty Th17-inducing strains isolated from human feces show epithelial-adhesive property
- Bacterial adhesion elicits a Th17-inducing gene-expression program in epithelium

Accession Numbers

GSE71734

Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells

Koji Atarashi,^{2,4,13} Takeshi Tanoue,^{2,13} Minoru Ando,^{1,13} Nobuhiko Kamada,^{5,13} Yuji Nagano,² Seiko Narushima,² Wataru Suda,^{4,7} Akemi Imaoka,¹ Hiromi Setoyama,¹ Takashi Nagamori,^{1,8} Eiji Ishikawa,¹ Tatsuchihiro Shima,¹ Taeko Hara,¹ Shoichi Kado,¹ Toshi Jinnohara,² Hiroshi Ohno,² Takashi Kondo,² Kiminori Toyooka,³ Eiichiro Watanabe,² Shin-ichiro Yokoyama,¹⁰ Shunji Tokoro,¹⁰ Hiroshi Mori,¹⁰ Yurika Noguchi,⁸ Hidetoshi Morita,^{8,12} Ivaylo I. Ivanov,⁹ Tsuyoshi Sugiyama,¹⁰ Gabriel Nuñez,⁶ J. Gray Camp,¹¹ Masahira Hattori,^{4,7} Yoshinori Umesaki,^{1,*} and Kenya Honda^{2,4,12,*}

¹Yakult Central Institute, 5-11 Izumi, Kunitachi, Tokyo 186-8650, Japan

²RIKEN Center for Integrative Medical Sciences (IMS)

³Center for Sustainable Resource Science (CSRS)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

⁴Department of Microbiology and Immunology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

⁵Division of Gastroenterology, Department of Internal Medicine

⁶Department of Pathology and Comprehensive Cancer Center

University of Michigan Medical School, Ann Arbor, MI 48109, USA

⁷Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

⁸School of Veterinary Medicine, Azabu University, 1-17-71 Fuchinobe, Sagami-cho, Kanagawa 252-5201, Japan

⁹Department of Microbiology and Immunology, Columbia University Medical Center, New York, NY 10032, USA

¹⁰Laboratory of Microbiology and Immunology, Department of Biopharmaceutical Sciences, Gifu Pharmaceutical University, 1-25-4 Daigaku-Nishi, Gifu 501-1196, Japan

¹¹Department of Evolutionary Genetics, Max Planck Institute of Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

¹²CREST, JST, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan

¹³Co-first author

*Correspondence: yoshinori-umesaki@yakult.co.jp (Y.U.), kenya@keio.jp (K.H.)

<http://dx.doi.org/10.1016/j.cell.2015.08.058>

SUMMARY

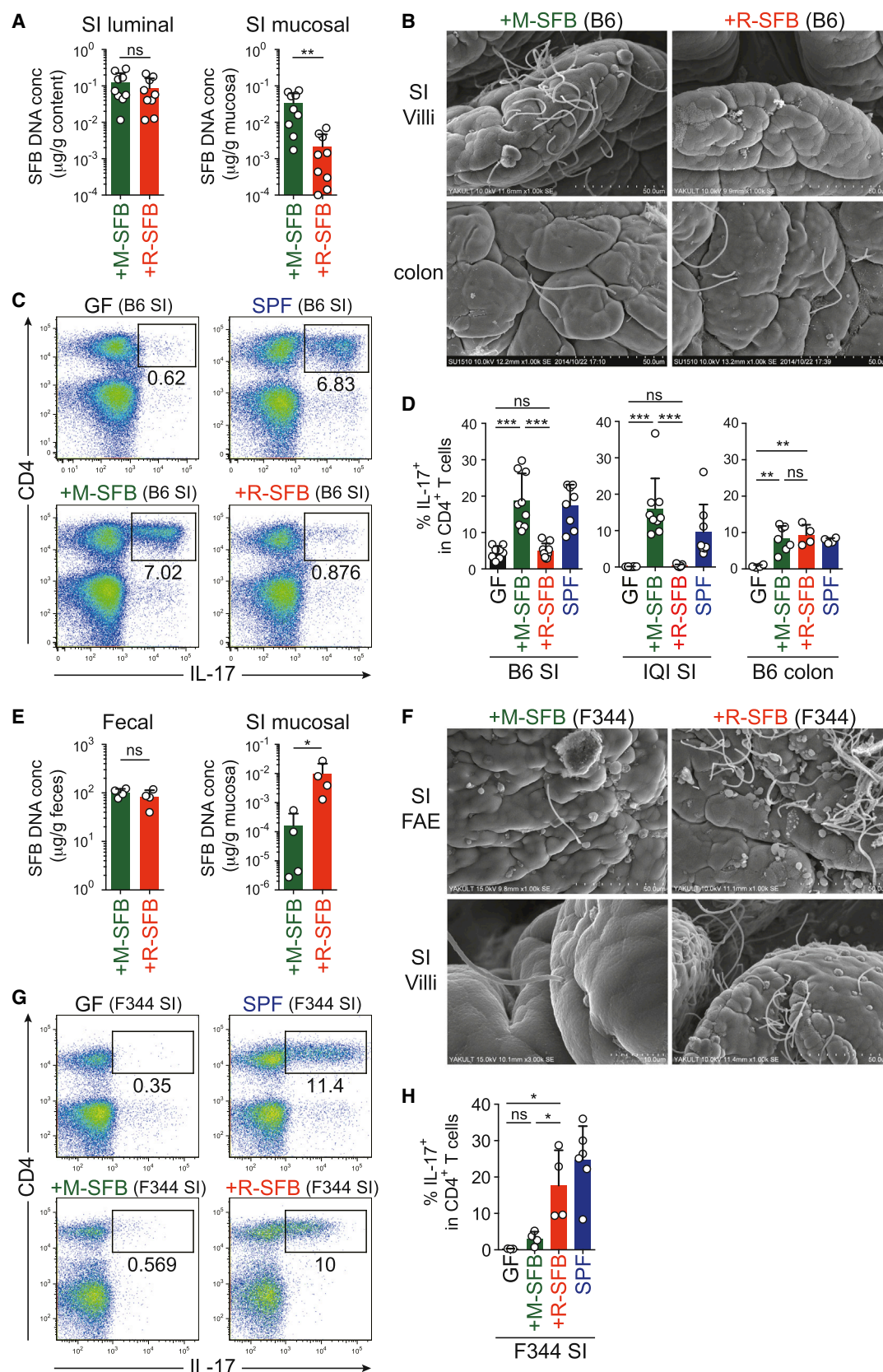
Intestinal Th17 cells are induced and accumulate in response to colonization with a subgroup of intestinal microbes such as segmented filamentous bacteria (SFB) and certain extracellular pathogens. Here, we show that adhesion of microbes to intestinal epithelial cells (ECs) is a critical cue for Th17 induction. Upon monocolonization of germ-free mice or rats with SFB indigenous to mice (M-SFB) or rats (R-SFB), M-SFB and R-SFB showed host-specific adhesion to small intestinal ECs, accompanied by host-specific induction of Th17 cells. *Citrobacter rodentium* and *Escherichia coli* O157 triggered similar Th17 responses, whereas adhesion-defective mutants of these microbes failed to do so. Moreover, a mixture of 20 bacterial strains, which were selected and isolated from fecal samples of a patient with ulcerative colitis on the basis of their ability to cause a robust induction of Th17 cells in the mouse colon, also exhibited EC-adhesive characteristics.

INTRODUCTION

The gut microbiota contributes to the constitutive development of Th17 cells in the intestinal lamina propria (LP) (Atarashi et al., 2008; Ivanov et al., 2008). Among commensals, segmented

filamentous bacteria (SFB) are one of the most potent inducers of Th17 cells, and monocolonization of mice with SFB causes abundant accumulation of Th17 cells in the small intestinal (SI) LP (Gaboriau-Routhiau et al., 2009; Ivanov et al., 2009). Recent reports have shown that most of the intestinal Th17 cells induced by SFB have T cell receptors (TCRs) that specifically recognize SFB antigens (Goto et al., 2014; Yang et al., 2014). However, since the SFB antigens themselves do not dictate Th17 differentiation (Yang et al., 2014), and microbiota-mediated Th17 cell development occurs independently of major innate immune receptors (Atarashi et al., 2008; Ivanov et al., 2009), SFB colonization must elicit unique signaling pathways in the intestine to generate a Th17-conducive environment.

SFB are spore-forming gram-positive bacteria with a segmented and filamentous morphology, and tight adhesion to SI epithelial cells (ECs) is a remarkable characteristic feature of these bacteria (Davis and Savage, 1974). SFB are widely distributed in vertebrates (Klaasen et al., 1993). In spite of the morphological similarities of SFB isolated from various hosts, their 16S rRNA gene sequences differ, and several reports suggest that SFB have undergone host species-specific selection and adaptation (Chung et al., 2012). The complete genomic sequences of SFB colonizing the mouse and rat intestines, referred to as M-SFB and R-SFB, respectively, were determined. Although the overall genomic organization of M-SFB and R-SFB are similar, 5%–10% of the genes are specific to each strain, and the amino acid sequence identity between orthologous gene pairs is on average 80% (Prakash et al., 2011). Analysis



(legend on next page)

of differences between M-SFB and R-SFB may be useful to improve understanding of the effects of SFB on the immune system.

In addition to SFB colonization, infections with several extracellular pathogens such as *Candida albicans* and *Citrobacter rodentium* are known to induce Th17 cells (Conti and Gaffen, 2010; Mangan et al., 2006). Th17 cells induce the recruitment of neutrophils and activation of ECs, leading to enhanced clearance of extracellular pathogens in concert with other immune cells such as IgA-secreting plasma cells and group 3 innate lymphoid cells (ILC3s). The induction of Th17 cells by those pathogens has been postulated to be mediated by the local cytokine milieu produced by intestinal ECs and specific subsets of myeloid cells (Weaver et al., 2013). However, it remains unclear which features of these particular microbes specifically elicit Th17 versus other types of immune cell responses at intestinal mucosal sites.

Because SFB and *C. rodentium* commonly adhere to ECs, we hypothesized that adhesion-mediated activation of ECs plays a pivotal role in the induction of Th17 cells. Accordingly, we examined the ability of M-SFB, R-SFB, wild-type, and mutant strains of *C. rodentium* and enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 to adhere to ECs and induce Th17 cells. In addition, by combining gnotobiotic technique and anaerobic culturing of members of the intestinal microbiota from a patient with ulcerative colitis (UC), we isolated 20 strains based on their ability to induce Th17 cells in mice and examined EC-adhesive characteristics of these 20 Th17-inducing human strains. Our findings indicate that adhesion to ECs is a common mechanism used by intestinal microbes to activate host Th17 responses.

RESULTS

Host-Specific Adhesion to SI ECs and Th17 Induction by SFB

C57BL/6 (B6) or IqI germ-free (GF) mice were orally inoculated with R-SFB or M-SFB, and their intestinal colonization was monitored by qPCR analysis. The concentration of fecal and SI luminal R-SFB DNA quickly increased and reached a plateau within 1 week; the kinetics and levels were comparable to those of M-SFB (Figures 1A and S1A). Consistent with the qPCR results, Gram-stained smears of cecal luminal contents contained equivalent numbers of R-SFB and M-SFB with indistinguishable morphology (Figure S1B), indicating that R-SFB and M-SFB both colonize and grow robustly within the mouse intestinal lumen. In contrast, when SI mucosa-associated SFB DNA amounts were examined, we detected much lower levels of R-SFB than of M-SFB (Figure 1A). We also performed scanning electron microscopy (SEM) of washed SI mucosa to visualize EC-adhering SFB. Numerous M-SFB were observed adhering tightly to the

mouse SI epithelium. In contrast, we did not detect any R-SFB adhering to SI ECs (Figures 1B and S1C). R-SFB also failed to adhere to SI ECs in monocolonized *Rag1*^{-/-} mice or *IL-2Rγ*^{-/-}*Rag2*^{-/-} mice (Figures S1D and S1E), making it unlikely that the impaired adhesion of R-SFB to mouse SI ECs was due to a R-SFB-specific reaction by the mouse immune system. Although we detected some EC-adhering M-SFB and R-SFB in the mouse colon, these were less frequent than in the SI (Figure 1B).

We then investigated the effects of monocolonization of mice with M-SFB or R-SFB on the induction of Th17 cells. In agreement with previous reports (Gaboriau-Routhiau et al., 2009; Ivanov et al., 2009), a robust accumulation of Th17 cells was observed in the SI LP of B6 or IqI mice monocolonized with M-SFB. In contrast, R-SFB colonization had no significant effect on the number of Th17 cells (Figures 1C and 1D). Notably, both M-SFB and R-SFB induced the accumulation of Th17 cells in the colonic LP, although less efficiently than that in the SI LP following M-SFB colonization (Figure 1D), correlating with the EC adhesion efficacy of the SFB.

We next performed the reciprocal experiment by monocolonizing GF F344 rats with R-SFB or M-SFB. After oral inoculation of GF rats with R-SFB, high levels of R-SFB DNA were detected in feces and SI mucosal tissues by qPCR (Figure 1E). Examination of the rat SI mucosal surface by SEM revealed the frequent presence of R-SFB adhering tightly to the follicular-associated epithelium (FAE) and villous ECs (Figure 1F). In contrast, far fewer M-SFB associated with the rat SI ECs, although the amount of M-SFB and R-SFB DNA in the feces was comparable (Figures 1E and 1F). We then investigated the effects of R-SFB and M-SFB monocolonization on Th17 cells in rats. The frequency of Th17 cells in the SI LP was markedly increased after monocolonization of GF rats with R-SFB, whereas M-SFB-monocolonization induced only a slight increase in the frequency of these cells (Figures 1G and 1H). Taken together, adhesion of SFB to intestinal ECs is a host-specific event and strongly correlates with Th17 cell induction.

Induction of Antigen-Specific Th17 Cells by EC-Adhesive SFB

To examine whether EC adhesion is required for priming SFB antigen-specific Th17 cell differentiation, we generated mice co-colonized with M-SFB and R-SFB and assessed the antigen specificity of the SI LP Th17 cells. In the feces of co-colonized mice, similar amounts of M-SFB and R-SFB DNA were detected (Figure S2A). SI LP cells, which include T cells and antigen-presenting cells (APCs), were isolated from the co-colonized mice, and stimulated ex vivo either with PMA and ionomycin (P/I) or with autoclaved cecal content from GF, M-SFB-, or R-SFB-monocolonized mice. Cytokine expression was analyzed as a

Figure 1. EC Adhesion and Th17 Induction by SFB in Mice and Rats

B6 or IqI mice (A–D) and F344 rats (E–H) were monocolonized with M-SFB or R-SFB for 3 weeks.

(A and E) qPCR analysis for SFB DNA in feces, SI luminal contents, and mucosal tissue specimens.

(B and F) SEM images of epithelial surfaces in the SI and colon.

(C, D, G, and H) Th17 cell frequencies in SI and colon LP. Representative dot plots gated on total lymphocytes (C and G) and summarized data gated on CD4⁺ T cells (D and H) are shown.

Error bars represent SD. See also Figure S1.

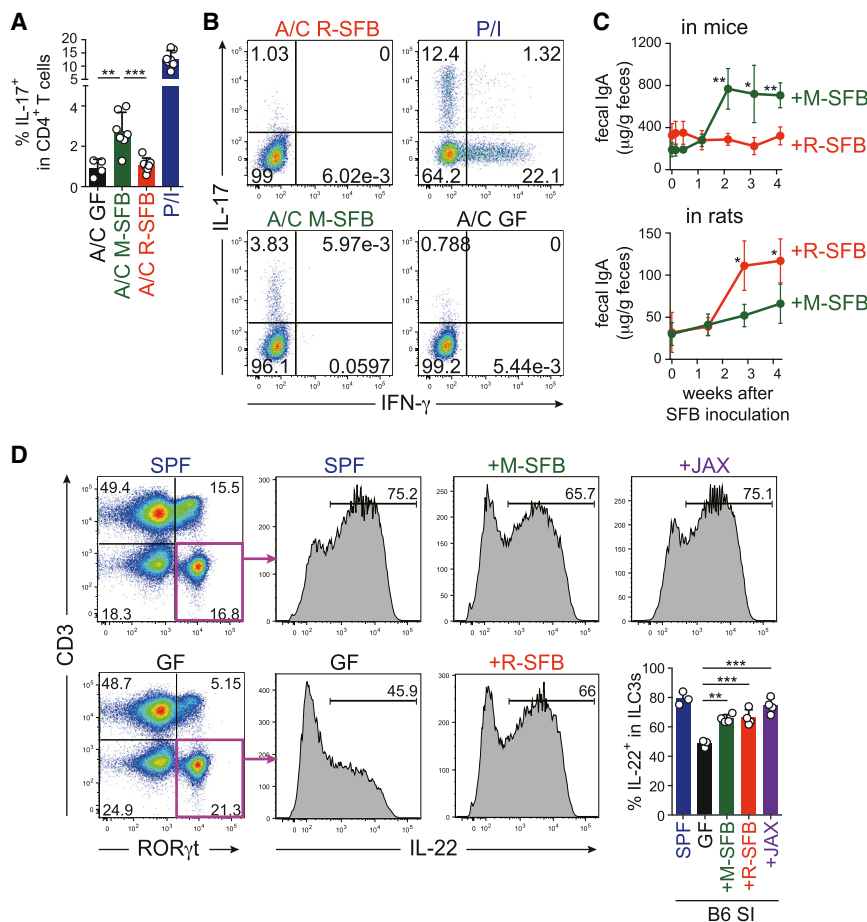


Figure 2. Effects of SFB Adhesion on Anti-gen-Specific Th17 Cells, IgA⁺ Cells, and ILC3s

(A and B) Cytokine responses of SI LP CD4⁺ T cells from M-SFB and R-SFB co-colonized B6 mice to the indicated ex vivo stimulation. A/C, autoclaved cecal contents; P/I, PMA/ionomycin.

(C) Fecal IgA levels of indicated monocolonized B6 mice and F344 rats (n = 5).

(D) The frequencies of ILC3s among CD90⁺ cells (left dot plots) and IL22⁺ cells among ILC3s (right histograms and bar graphs) in SI LP cells of the indicated mice.

Error bars represent SD. See also Figure S2.

Effects of EC-Adhering SFB on IgA⁺ and ILC3 Cells

SFB colonization has been shown to enhance IgA production (Talham et al., 1999; Umesaki et al., 1999). Consistent with these reports, monocolonization of B6 mice with M-SFB induced a significant increase in the levels of fecal IgA and frequencies of IgA⁺CD138⁺ plasma cells in the SI LP (Figures 2C and S2C). In contrast, R-SFB monocolonization in mice resulted in a weak IgA response (Figures 2C and S2C). We also examined IgA production in SFB monocolonized rats. Fecal IgA levels were substantially increased after monocolonization of GF rats with R-SFB, whereas M-SFB-monocolonization induced only a modest increase (Figure 2C). These findings suggest that luminal colonization is insufficient for the full induction of IgA⁺ cells, but instead an interaction between SFB and ECs is required, which is similar to Th17 cells.

Next, we examined the effects of SFB colonization on ILC3s, which share many features with Th17 cells (Spits et al., 2013). Thy1⁺CD3⁺RORγt⁺ cells corresponding to ILC3s were much more abundant in the SI LP of SPF B6 mice than in the spleen (Figure S2D), and ~70% of these cells became positive for IL-22 after ex vivo stimulation with IL-23 (Figure 2D). Although the abundance of SI LP ILC3s in GF versus SPF mice did not differ significantly, the frequency of IL-22-proficient (IL-22⁺) cells among ILC3s was dramatically decreased in GF mice (Figure 2D). Monocolonization of mice with M-SFB resulted in a significant increase in the frequency of IL-22⁺ SI ILC3s, although the level did not reach that observed in SPF mice (Figure 2D). The increase in IL-22⁺ cells upon monocolonization with M-SFB was observed for both NKp46⁺ and NKp46[−] ILC3 populations (Figure S2E). Interestingly, R-SFB induced a significant increase in the frequency of IL-22⁺ cells with a magnitude similar to that of M-SFB (Figure 2D). The induction of IL-22⁺ ILC3s was also observed in mice inoculated with feces from Jackson Laboratory (JAX) SPF mice, which are known to be devoid of SFB (Ivanov et al., 2009) (Figure 2D), indicating that IL-22⁺ ILC3s can be induced by non-adhering

readout for TCR activation. Although less efficient than P/I, the M-SFB antigens evoked IL-17 expression in a significant number of cells (Figures 2A and 2B). In contrast, the R-SFB antigens elicited no response. Therefore, a substantial fraction of Th17 cells that accumulated in the SI LP of co-colonized mice had TCRs specific to the EC-adhering M-SFB, but not to the non-adhering R-SFB. These results suggest that antigens from EC-adhering SFB are taken up and presented by APCs much more efficiently than those of non-adhering SFB.

Recently, two major protein antigens responsible for M-SFB-mediated SI LP Th17 cell induction were identified (SFBNYU_003340 and SFBNYU_004990) (Yang et al., 2014). We analyzed the R-SFB genome and identified the gene encoding the orthologous protein corresponding to SFBNYU_003340, but could not identify an SFBNYU_004990 ortholog. The minimal epitope within M-SFB SFBNYU_003340 was mapped as QFSGAVPNK (Yang et al., 2014), and the corresponding epitope region in R-SFB was found to be QFNGQVPNN. The M-SFB epitope peptide efficiently induced a Th17 response when added to the ex vivo cultures of SI LP cells isolated from specific pathogen-free (SPF) mice (Figure S2B). In contrast, the same SI LP cells had no reactivity to the R-SFB peptide QFNGQVPNN. Therefore, peptides serving as dominant antigens for SI LP Th17 cell induction in mice are specifically expressed by M-SFB.

SFB and even by non-SFB commensals, which is in contrast to Th17 and IgA⁺ cells.

SI EC Activation by Adhering SFB

We next examined the influence of SFB adhesion on SI EC gene expression profiles by RNA sequencing (RNA-seq). The expression of several genes was highly upregulated in SI ECs of M-SFB-colonized mice compared to GF and R-SFB-colonized mice (Figures 3A and S3A). The upregulated genes included three isoforms of serum amyloid A (SAA), regenerating islet-derived protein 3 β (Reg3 β) and Reg3 γ and nitric oxide synthase 2 (Nos2). The expression differences of selected genes were confirmed by qPCR analysis (Figures 3B and S3B). Elevated levels of SAA1 protein in SI ECs of M-SFB-monocolonized mice were clearly visualized by immunofluorescence microscopy (Figure 3C). Recombinant SAA1 markedly enhanced the in vitro differentiation of naive CD4⁺ T cells into Th17 cells mediated by CD11c⁺ cells, IL-6, and transforming growth factor β (TGF- β) (Figure 3D). This SAA1-mediated enhancement was severely attenuated in the absence of CD11c⁺ cells or of IL-6 and TGF- β , or in the presence of an anti-IL-1 receptor 1 (IL-1R1) blocking antibody (Figures 3D and 3E). In fact, SAA1 elevated the expression level of IL-1 β mRNA in splenic CD11c⁺ cells (Figure 3F). Therefore, the current data extend previous studies reporting that SAAs are induced in SI ECs by colonization with SFB (Ivanov et al., 2009) and suggest that SAA induction requires SFB adhesion to ECs. Furthermore, the SAAs produced by this mechanism likely condition neighboring CD11c⁺ myeloid cells to produce IL-1 β and probably other factors that act with IL-6 and TGF- β to enhance Th17 cell differentiation.

In addition to the increase in SAAs, mRNA levels of the reactive oxygen species (ROS)-generating enzyme dual oxidase 2 (Duox2) and of its maturation factor Duoxa2 were also highly upregulated in SI ECs in M-SFB-monocolonized mice (Figures 3A and 3B). Treatment of M-SFB-monocolonized mice with a ROS scavenger, N-acetyl-L-cysteine (NAC), in the drinking water significantly limited Th17 cell induction (Figure 3G) without affecting adhesion of M-SFB to SI ECs (Figure S3C). Therefore, EC-adhering SFB affect Th17 cell accumulation via integration of a number of different molecular mechanisms including production of SAAs and ROS by ECs.

Toward an understanding of the molecular basis of adherent SFB-mediated SAA induction, we explored the transcriptional regulatory elements of the *Saa1* gene locus. We identified two DNase I hypersensitive sites downstream of the *Saa1* gene, which were uniquely active in the SI and distinguished by acetylated histone 3 lysine 27 (H3K27ac) active enhancer marks (Shen et al., 2012) (Figure 3H). Investigation of publicly available datasets of chromatin immunoprecipitation sequencing (ChIP-seq) for transcription factor (TF) binding (Camp et al., 2014) suggested interactions of CCAAT-enhancer-binding protein (C/EBP) with both regulatory regions of *Saa1* (Figure 3H). Our SI EC RNA-seq data and qPCR analysis showed that C/EBP δ was specifically upregulated in SI ECs of M-SFB-monocolonized mice (Figures 3I and S3B), suggesting that it may be a key TF regulating SAA1 expression. In line with this finding, we observed co-induction of SAA1 and C/EBP δ expression in a mouse SI EC line (aMos7) co-cultured with M-SFB in vitro (Figure 3J).

Induction of SAA1 and C/EBP δ expression was observed only when ECs were in direct contact with M-SFB (Figure 3J); under these conditions M-SFB adhesion to the aMos7 cells was confirmed by fluorescence in situ hybridization (FISH) and SEM (Figures S3D and S3E). The induction of SAA1 and C/EBP δ expression by M-SFB was further enhanced by the addition of F-actin inhibitory drugs, latrunculin A, or swinholide A (Figure S3F). Therefore, actin reorganization in ECs induced by SFB adhesion may lead to the elevation of C/EBP δ expression, which then contributes to SAA1 expression.

Influence of Mouse Genetic Background on SFB-Mediated Th17 Cell Induction

All of the above experiments were performed using B6 or IQR mice. We next examined the influence of mouse genetic background on intestinal Th17 cell induction. The frequency of Th17 cells in SPF C57BL/10 and B10.D2 mice was similar to that in B6 and IQR (ICR) mice (Figure 4A). The only exception was BALB/c mice, which had a much lower frequency of Th17 cells even under SPF conditions or after monocolonization with M-SFB (Figures 4A, 4B, and S4A). It is noteworthy that BALB/c mice had substantial numbers of ROR γ t⁺ single-positive cells but had severe reduction of ROR γ t⁺IL-17⁺ double-positive cells (Figure S4A). M-SFB, but not R-SFB, were found to adhere to SI EC surfaces in BALB/c mice, similar to the results in B6 mice (Figures 4C and S4B) and splenic naive CD4⁺ T cells from BALB/c mice gave rise to a normal frequency of ROR γ t⁺IL-17⁺ cells when differentiated under Th17 conditions (Figure S4C), suggesting that SFB adhesion to SI ECs occurs normally and that CD4⁺ T cells do not have an intrinsic differentiation defect, but instead that EC activation and/or subsequent LP cell signaling required for ROR γ t⁺IL-17⁺ Th17 induction may be defective in BALB/c mice.

Among Th17-conducting cytokines, microbiota-induced constitutive and high-level expression of IL-1 β was observed in SI LP CD11c⁺ cells of B6 mice (Figure 4D), which is in line with previous reports (Mortha et al., 2014; Shaw et al., 2012). IL-1 β expression was significantly reduced in M-SFB-monocolonized and SPF BALB/c mice, compared with that of B6 mice (Figure 4E). Besides the reduced expression of IL-1 β in CD11c⁺ cells, expression of SAA1/2 was significantly decreased in SI ECs of BALB/c mice (Figures 4F and S4D). Injection of exogenous IL-1 β into SPF BALB/c mice elevated expression of SAAs in SI ECs (Figure 4G) and this was accompanied by an increase in SI LP ROR γ t⁺IL-17⁺ Th17 cells (Figures 4J and S4E), suggesting that insufficient IL-1 β production by LP CD11c⁺ cells is responsible, at least in part, for the reduction in SAA expression and Th17 cell accumulation in BALB/c mice. Consistent with this interpretation, B6 *Il1r1*^{-/-} mice phenocopied BALB/c mice, displaying decreased expression of SAAs in SI ECs (Figure 4H) and a significant reduction in the number of SI LP ROR γ t⁺IL-17⁺ Th17 cells (Figures 4K and S4F). It is noteworthy that IL-1 β administration induced Th17 cell accumulation in M-SFB-colonized mice, but not in R-SFB-colonized or GF BALB/c mice, revealing a requirement for the presence of EC-adhering bacteria for the IL-1 β effect (Figures 4J and S4E).

BALB/c mice had reduced frequencies of IL-22⁺ cells among ILC3s compared with B6 mice under both SPF and

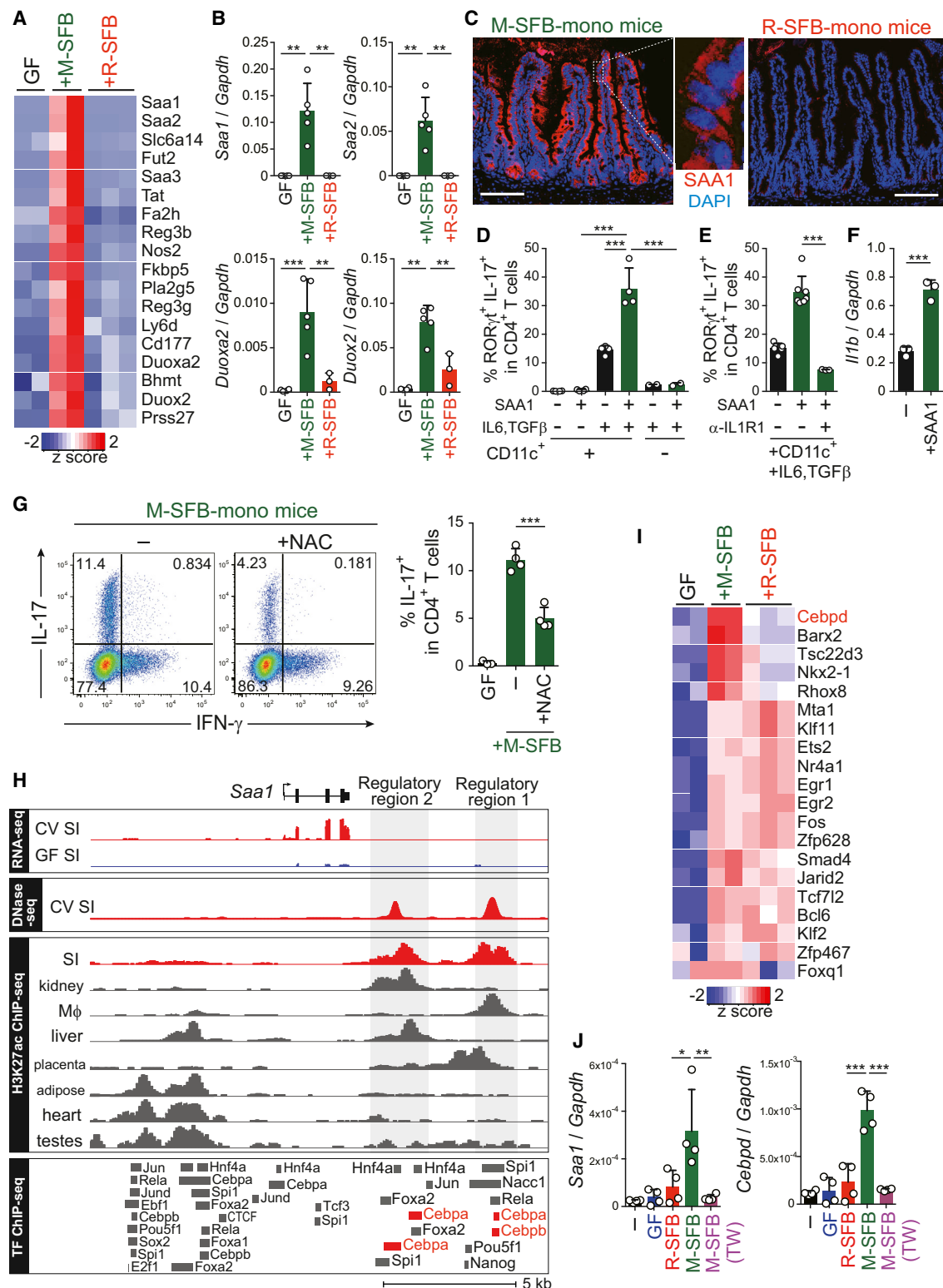


Figure 3. Adhesion-Mediated EC Activation

(A) Heatmap showing the relative abundance for gene transcripts upregulated in SI ECs of M-SFB-mono versus GF and R-SFB-mono IQI mice. Each column represents a single mouse.

(B) qPCR for the selected genes relative to *Gapdh* in SI ECs from the indicated mice.

(legend continued on next page)

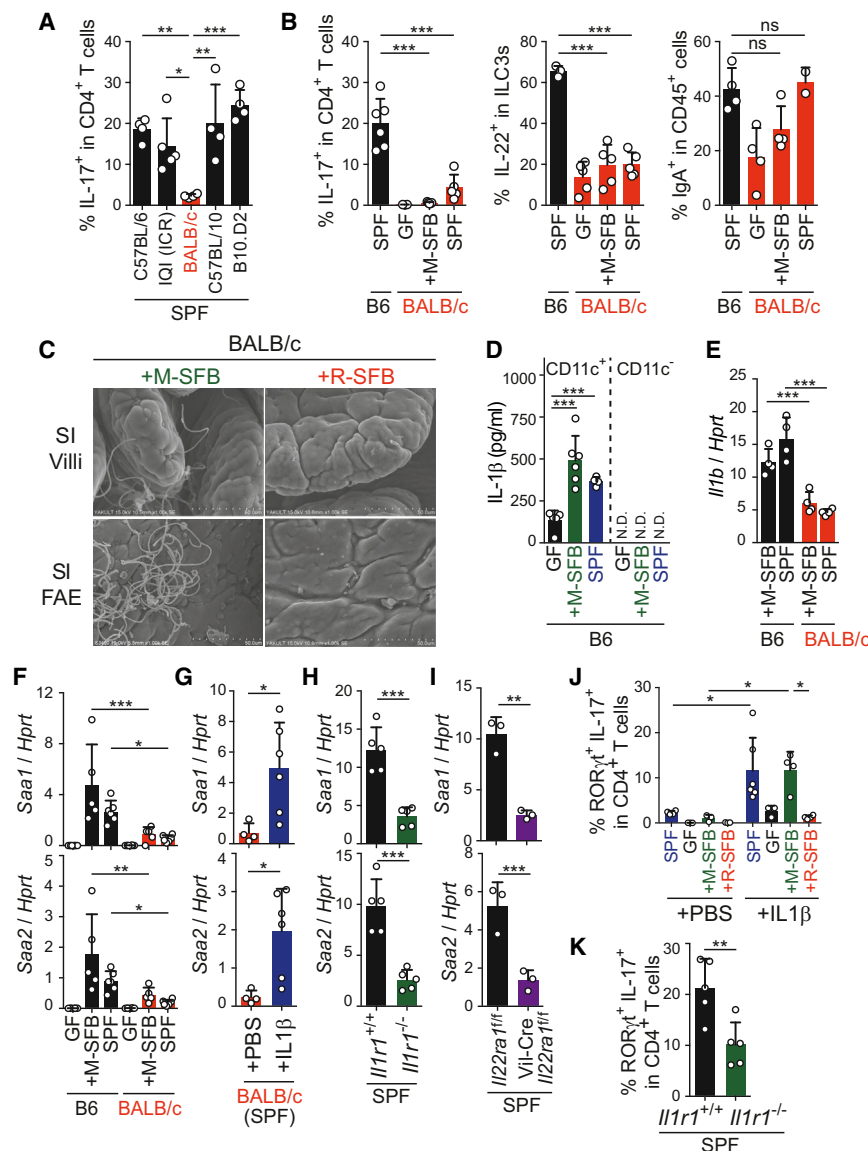


Figure 4. Influences of Mouse Genetic Background

(A and B) The percentages of Th17, IL-22⁺ ILC3, and CD138⁺IgA⁺ cells in SI LP of the indicated mice.

(C) SEM images of SI epithelial surface of the indicated BALB/c mice.

(D and E) IL-1β protein production (D) and mRNA expression (E) from SI LP CD11c⁺ and CD11c⁻ cells from the indicated mice.

(F–I) *Saa1/2* mRNA expression in SI ECs from the indicated mice.

(J and K) The percentage of RORγt⁺IL-17⁺ cells among SI LP CD4⁺ T cells in IL-1β-injected BALB/c mice (J) or *Il1r1*^{-/-} B6 mice (K).

Error bars represent SD. See also Figure S4.

sion of SAAs by SI ECs (Figure 4I), suggesting that IL-22 from ILC3s, together with IL-1β from CD11c⁺ cells, act on SI ECs to potentiate their expression of SAA. Vil-Cre crossed with *Il22ra1*^{fl/fl} mice showed a tendency toward reduction of RORγt⁺IL-17⁺ Th17 cells in the SI LP (Figure S4H). Therefore, adhesive SFB may create a complex circuitry of interactions between ECs, DCs and ILC3s mediated by SAA, IL-1β, IL-22, ROS, and probably other factors, to amplify and maintain signaling for constitutive accumulation of Th17 cells (Figure S4I).

Extracellular Pathogens Induce Identical Th17 Responses through EC Adhesion

To further test the link between bacterial EC adhesion and Th17 induction, we examined *Citrobacter rodentium*, an attaching/effacing enteropathogen known to induce Th17 cells in mice (Mangan et al., 2006). B6 GF mice were mono-colonized with wild-type or a mutant

M-SFB-mono-colonized conditions, whereas the frequency of IgA⁺ cells was normal (Figures 4B and S4G). An EC-specific deficiency of the IL-22 receptor α1 gene (Vil-Cre crossed with *Il22ra1*^{fl/fl}) on an SPF B6 background resulted in reduced expres-

C. rodentium strain lacking the *eae* gene, which encodes Intimin, a protein essential for EC adhesion (Kamada et al., 2012). The Δ*eae* and wild-type *C. rodentium* expanded equivalently in the intestinal lumen, as demonstrated by recovery of similar

(C) Immunostaining of SIs from the corresponding mice for SAA1 (red) and DAPI (blue). Scale bar, 100 μm.

(D and E) The percentage of RORγt⁺IL-17⁺ cells after culture of splenic naive CD4 T cells in the presence of various combinations of splenic CD11c⁺ cells, IL-6, TGF-β, SAA1, and/or anti-IL-1R1 antibody.

(F) *Il1b* expression in splenic CD11c⁺ cells from B6 SPF mice stimulated with recombinant SAA1.

(G) SI LP Th17 cell frequencies of IQ1 M-SFB-mono mice either untreated (–) or treated with NAC.

(H) RNA-seq and DNase-seq data from SI ECs of GF and conventional (CV) mice at the region surrounding the *Saa1* locus. H3K27ac ChIP-seq data for SI and various other tissues obtained from (Shen et al., 2012). ChIP-seq datasets for TFs assayed in various cell and tissue types. SI EC regulatory regions are highlighted in gray.

(I) The relative abundance for TF genes that are >2-fold different between M-SFB-mono mice versus GF or R-SFB-mono mice.

(J) SAA1 and C/EBPβ genes expression in aMoS7 cells cultured with cecal content from GF mice or a Percoll-enriched M-SFB or R-SFB either in contact or separated by a Transwell membrane (TW).

Error bars represent SD. See also Figure S3.

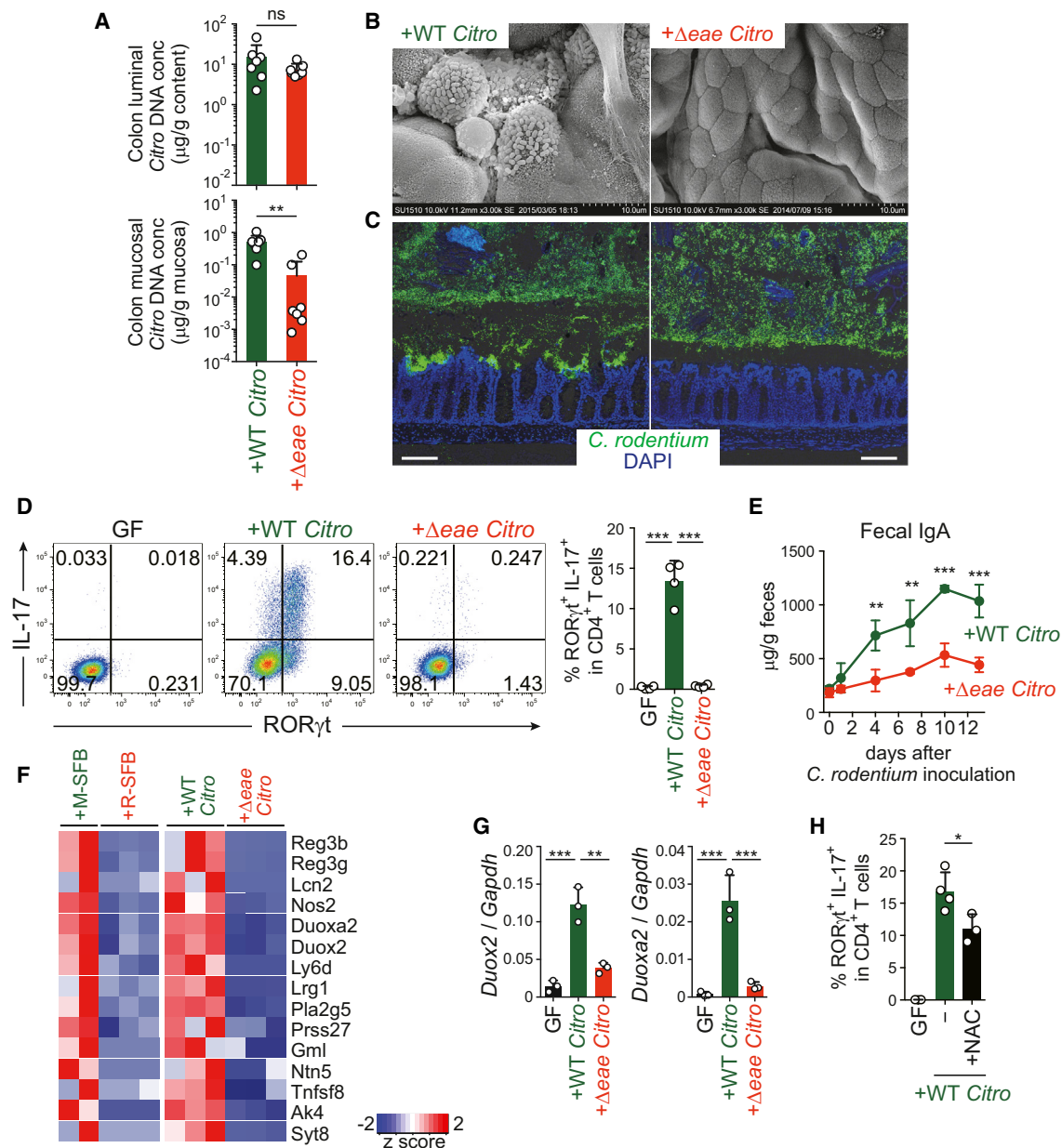


Figure 5. EC Adhesion-Mediated Th17 Induction by *C. rodentium*

IQI mice were monocolonized with wild-type (WT) or Δeae mutant of *C. rodentium* for 5 days.

(A) DNA of *C. rodentium* in colonic luminal contents and mucosal tissues.

(B) SEM images of colonic villi of the indicated mice.

(C) *C. rodentium* localization visualized by O antigen antisera (green) and DAPI (blue) staining. Scale bar, 100 μm .

(D) The percentage of $\text{ROR}\gamma\text{t}^+\text{IL-17}^+$ colonic CD4^+ T cells.

(E) Fecal IgA levels of the indicated mice (n = 4).

(F) The relative abundance of gene transcripts that were commonly upregulated in M-SFB-mono mice versus R-SFB-mono mice and WT *C. rodentium*-mono mice versus Δeae *C. rodentium*-mono mice.

(G) *Duox2* and *Duoxa2* mRNA expression in colonic ECs of the indicated mice.

(H) Th17 cell frequencies of WT *C. rodentium*-mono IQI mice either untreated (–) or treated with NAC.

Error bars represent SD. See also Figure S5.

amounts of bacterial DNA from the colon luminal contents of both groups of mice after 5 days of monocolonization (Figure 5A). SEM and immunofluorescence microscopy revealed that the

wild-type strain had a strong tendency to adhere to colonic ECs, whereas the Δeae strain did not adhere to ECs and the inner mucus layer of the colon remained intact (Figures 5B and

5C). These results were further confirmed by qPCR quantification of mucosa-associated *C. rodentium* (Figure 5A). Colonization with wild-type *C. rodentium* resulted in potent induction of Th17 cells in the colon, whereas Δeae *C. rodentium* elicited a far weaker Th17 response (Figure 5D). Colonization with wild-type *C. rodentium* induced a greater increase in fecal IgA levels than the mutant strain (Figure 5E). Therefore, similarly to commensal SFB, pathogenic *C. rodentium* induces Th17 and IgA responses, most likely through adhesion to ECs. Of note, during the early phase (5 days after infection), when the epithelial layer was relatively intact, infection with wild-type *C. rodentium* caused a more conspicuous increase in Th17 cells than IFN- γ -producing Th1 cells. However, in the late phase (14 days after infection), when the epithelial layer was severely disrupted and bacterial invasion had occurred, there was a massive increase in Th1 cells rather than Th17 cells (Figures S5A and S5B), supporting the requirement of intact ECs for Th17 induction in response to *C. rodentium*, without which the same microbe would elicit a different type of immune response. We also examined IL-22⁺ ILC3 induction by *C. rodentium* colonization. The Δeae strain induced a robust increase in the frequency of colonic IL-22⁺ ILC3 cells comparable to the response observed in mice colonized with the wild-type strain (Figure S5C), indicating that immune stimulatory activity is not completely abrogated in the Δeae strain, and the induction of IL-22⁺ ILC3 cells occurs independently of bacterial EC-adhesion.

Next, we performed gene expression analysis by RNA-seq of colon ECs from mice monocolonized with the Δeae or wild-type *C. rodentium* strain. In comparison to the Δeae strain, the wild-type strain specifically induced the expression of genes, such as *Duox2*, *Duoxa2*, *RegIII β* , *RegIII γ* , and *Nos2*, which overlap with genes induced in SI ECs by colonization with EC-adherent SFB (Figures 5F, 5G, S5D, and S5E), although the expression of SAAs in the colon was not significantly enhanced by *C. rodentium* monocolonization (Figure S5D). Treatment of wild-type *C. rodentium*-monocolonized mice with NAC in the drinking water significantly limited Th17 cell induction (Figure 5H), without affecting adhesion of *C. rodentium* to ECs (Figure S5F). Therefore, EC adhesion and subsequent ROS production is a potential mechanism underlying *C. rodentium*-mediated Th17 cell induction.

We next examined EHEC O157:H7, another EC-adherent bacterium that can cause hemolytic uremic syndrome in humans and is a global public health concern (Croxen and Finlay, 2010). Because the wild-type O157 strain was lethal in GF mice shortly after the monocolonization (data not shown), we used an O157 strain deficient in both Shiga toxin 1 and 2 genes ($\Delta stx1\Delta stx2$) (Yokoyama et al., 2001). This double mutant strain expanded in the intestine upon monocolonization without killing the mice and adhered to colonic EC surfaces (Figures 6A and 6B). The adhesion was accompanied by induction of colonic Th17 cells (Figure 6C). We subsequently constructed a $\Delta stx1\Delta stx2\Delta eae$ triple mutant strain, which expanded comparably to the $\Delta stx1\Delta stx2$ double mutant strain in the intestinal lumen but did not adhere to colonic ECs (Figures 6A and 6B), and this was accompanied by loss of the ability to induce Th17 cells (Figure 6C).

Candida albicans is also an EC-adhesive microbe (Figure S6) and was found to increase the frequency of Th17 cells in the

colon LP of monocolonized mice (Figure 6D). In contrast, monocolonization of mice with *Listeria monocytogenes* mainly induced Th1 cells, rather than Th17 cells (Figure 6E). The mice colonized with *L. monocytogenes* displayed disrupted structure of colon epithelial surface, whereas adhering bacteria were rarely observed (Figure 6F).

Th17-Inducing Human Microbiota

Finally, we attempted to identify members of the human gut microbiota that could exert immunological effects equivalent to those of SFB. To this end, we collected human fecal samples from patients with UC (UC4-2 and UC5-1) and healthy donors (H11, H17, H23) and inoculated them into GF IQI mice. All groups of mice exhibited a significant increase in Th17 cells in the colon (Figure 7A). The Th17 cell induction by H23 microbiota was suppressed by treatment with ampicillin (Amp) or other antibiotics (Figure 7B). In contrast, the UC5-1 microbiota-mediated Th17 cell induction was further enhanced when the mice were given Amp in the drinking water, although reduced by metronidazole (MNZ) or vancomycin (VCM) treatment (Figure 7B). Therefore, different subsets of bacterial species were likely to be responsible for the induction of Th17 cells in mice inoculated with H23 versus UC5-1 samples. Cecal microbiota analysis by 16S rRNA gene sequencing revealed several operational taxonomic units (OTUs) that were enriched in mice inoculated with UC5-1 sample treated with Amp and correlated with the number of colonic Th17 cells (Figure 7C, marked in red). To isolate Th17-inducing bacterial species, we anaerobically cultured cecal contents from Amp-treated UC5-1 mice using various culture media and picked 192 colonies. BLAST searches of 16S rRNA gene sequences revealed that the isolates contained 20 strains that broadly covered the bacterial species colonizing Amp-treated UC5-1 mice (Figure 7C, marked in green). Their genomes were sequenced and a phylogenetic comparison was carried out using 19 ribosomal protein genes predicted from the assembled draft genome of each strain. This analysis revealed that the 20 strains included diverse species, such as *Clostridium*, *Bifidobacterium*, *Ruminococcus*, and *Bacteroides* (Figures 7C and S7A; Table S1).

To investigate whether the isolated 20 strains have the ability to induce Th17 cells, we introduced a mixture of 20 strains into GF IQI mice. In mice inoculated with the 20 strains, we observed a robust accumulation of Th17 cells in the colon (Figures 7D and 7E). The 20-strain mix was also effective in Th17 induction in rats (Figure 7F). Similarly to SFB, the 20 strains enhanced expression of *Nos2* and *Duoxa2* in ECs and IgA and ILC3 responses (Figure 7G), whereas no significant effect was observed for Th1 cells in the colon (Figures 7D and 7E). To examine whether these 20 Th17-inducing strains contained EC-adherent bacteria, we performed SEM and FISH. As a control, we examined mice colonized with 17 strains of human-derived *Clostridia* that have the capacity to induce regulatory T cells (Tregs) (Atarashi et al., 2013). The 17 Treg-inducing strains of *Clostridia* were not adhering and disappeared after washing. In contrast, we observed numerous bacteria adhering to the colonic ECs in mice and rats colonized with the 20 Th17-inducing strains (Figures 7H and 7I), which is in line with our hypothesis that EC adhesion is an important signal for Th17 cell induction.

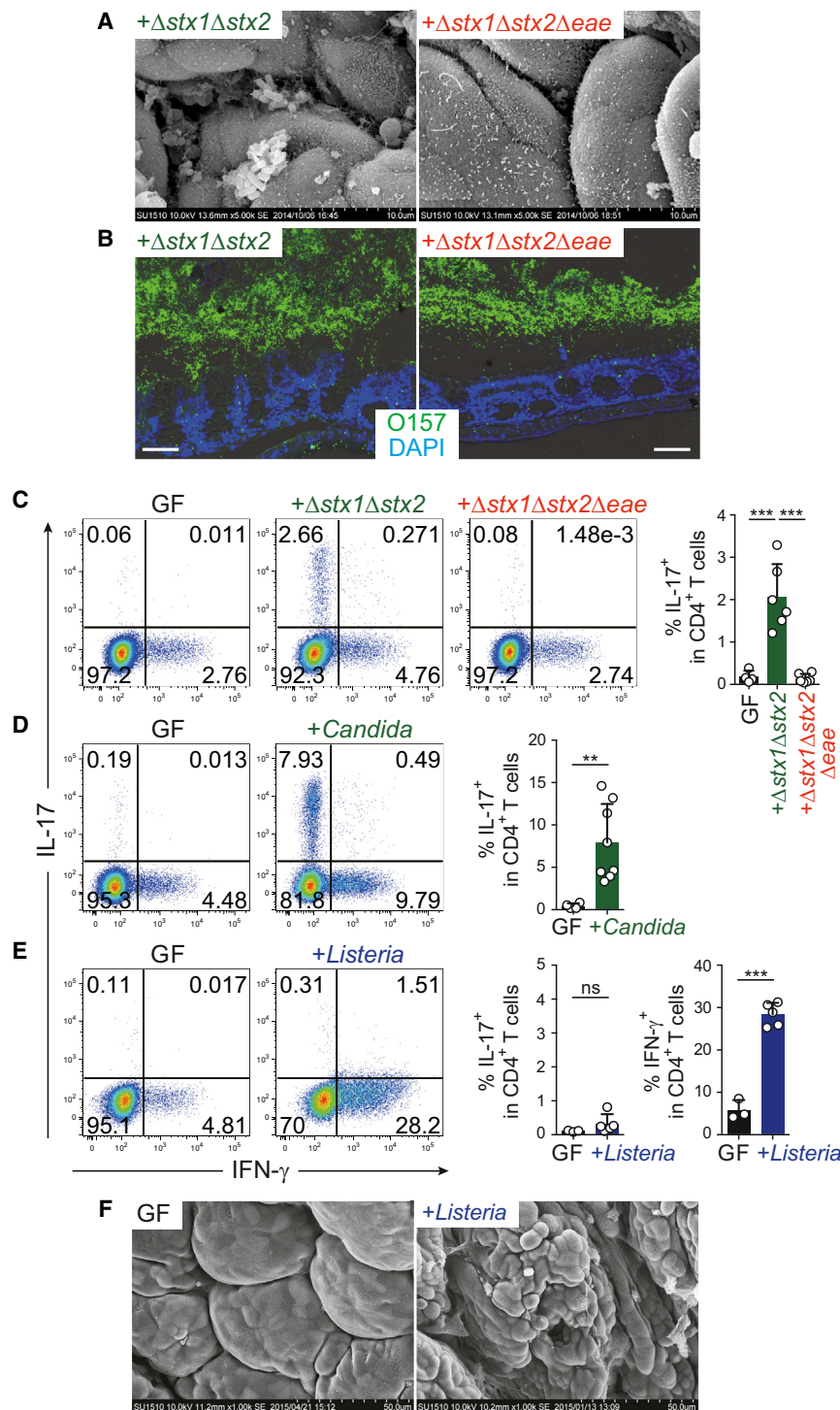


Figure 6. EC Adhesion-Mediated Th17 Induction by Extracellular Pathogens

(A) SEM images of colonic villi from IQI GF mice monocolonized with $\Delta stx1\Delta stx2$ or $\Delta stx1\Delta stx2\Delta eae$ EHEC O157:H7 for 3 weeks.

(B) O antigen antisera (green) and DAPI (blue) staining for EHEC O157 visualization. Scale bar, 50 μm .

(C–E) Th17 and Th1 cell proportions in the colonic LP CD4⁺ T cells from IQI mice inoculated with EHEC O157:H7 (C), *C. albicans* (D), or *L. monocytogenes* (E).

(F) SEM images of the colon of IQI GF mice and mice infected orally with *L. monocytogenes*. Error bars represent SD. See also Figure S6.

cells in the mouse and rat intestine, and these 20 strains displayed similar EC adhesive characteristics. Therefore, the intestinal immune system mounts Th17 cell responses through recognition of a physical interaction with the microbes, rather than by recognition of released microbial components or soluble metabolites. EC adhesion by microbes was also associated with increases in intestinal IgA⁺ cells. Consistent with our observation, it has been shown that potentially colitogenic bacteria that can invade the inner mucus layer and colonize near the epithelium induce high-affinity IgA responses and become highly coated with IgA (Palm et al., 2014). In contrast to Th17 cells and IgA⁺ cells, IL-22⁺ ILC3s can be generated by nonadhering commensals. Therefore, Th17 and IgA⁺ cells control EC-adhering microbes, whereas IL-22⁺ ILC3s may play a more general role in regulation of gut microbiota. The vast majority of gut microbiota are physically separated by the mucus layer (Johansson et al., 2008; Vaishnava et al., 2011), whereas microbes that cross the mucus barrier and attach to the epithelial lining are in many cases pathogenic (SFB and the 20 human strains are exceptional in that they can cross the mucus layer and attach to ECs, but do not penetrate the EC layer). Therefore, it seems reasonable to propose that the host immune system has evolved

to discriminate the biogeographical distributions of intestinal microbes.

Several reports have demonstrated that adhering SFB induce actin reorganization in SI ECs (Davis and Savage, 1974; Jepson et al., 1993). In this context, morphologically similar cytoplasmic alterations in intestinal ECs have been described in infections

DISCUSSION

In the present study, we demonstrate that SFB, *C. rodentium* and EHEC O157 promote the induction of intestinal Th17 cells through adhesion to intestinal ECs. Moreover, we isolated 20 bacterial strains from human intestine that can induce Th17

with EHEC O157 and *C. rodentium* (Croxen and Finlay, 2010). It is noteworthy that F-actin inhibition resulted in potentiated C/EBP δ and SAA1 expression in aMoS7 cells cocultured with M-SFB (Figure S3F). It is interesting to note that SFB have putative ADP-ribosyl transferases that have sequence similarities to the *Clostridium perfringens* iota toxin, which ribosylates G-actin and thereby inhibits its polymerization (Pamp et al., 2012). Furthermore, because ROS suppresses the activity of certain Rho GTPase family members (Stanley et al., 2014), ROS produced by the Duox2/Duoxa2 system in ECs may, in addition or instead, affect actin reorganization. In any case, there may be a link between actin cytoskeletal modulation by adherent SFB and subsequent gene induction in ECs, possibly through mechanosensing mechanisms (Dupont et al., 2011).

Since recombinant SAA1 activated CD11c⁺ cells to upregulate IL-1 β expression (Figure 3F), and recombinant IL-1 β activated ECs to potentiate expression of SAAs (Figure 4G), our results suggest a model in which the initial increase in SAAs in SI ECs caused by SFB adhesion is then augmented via an amplification loop of SAA1 and IL-1 β between SI ECs and CD11c⁺ myeloid cells (Figure S4I). This amplification processes may be required for the maintenance of high level expression of SAAs in SI ECs and IL-1 β in myeloid cells, which causes constitutive accumulation of Th17 cells. In an accompanying article published in this issue of Cell, Sano et al. (2015) show that SAAs can directly act on ROR γ t⁺ T cells to induce IL-17 expression. Derebe et al. (2014) showed that SAAs act as a transporter of retinol. Therefore, SAAs may exert their effects through multiple mechanisms. IL-22 derived from ILC3s also contributes to this loop via potentiation of SAA expression by SI ECs. IL-22 alone cannot sufficiently activate this amplification loop, as was shown that IL-22⁺ ILC3s induced by luminal R-SFB cannot induce Th17 cells (Figure 2D).

A substantial fraction of Th17 cells specifically recognized EC-adhesive M-SFB but not non-adhering R-SFB (Figures 2A and 2B). The simplest explanation for this observation is that antigens may be taken up by host cells only when SFB adhere to ECs. This process likely involves bacterial sampling by dendritic cells (DCs). In this context, recognition and phagocytosis of *C. rodentium*-infected apoptotic ECs by LP DCs were shown to preferentially trigger Th17 cell differentiation (Torchinsky et al., 2009). SFB and other EC-adhesive microbes may be engulfed by LP DCs together with their adhering apoptotic ECs, processed and loaded onto MHC class II molecules. SFB are innocuous members of the microbiota; however, SFB can be pathogenic, depending on the host genotype. For example, SFB have been implicated in T cell-mediated autoimmune diseases such as arthritis (Wu et al., 2010). In this context, we observed a striking genotype-specific difference upon monocolonization with EC-adhesive SFB in the induction of Th17 cells, with BALB/c mice having fewer Th17 cells in comparison to B6 and IqI mice. Our observation in BALB/c mice provides one remarkable example showing that the combination of genetic background and composition of the gut microbiota affects T cell status. Genotype-specific differences in the differentiation of T cells in response to a component of the gut microbiota may explain the phenotypic variations in autoimmune disease susceptibility.

Different from the host-specific Th17 induction by SFB, the 20 bacterial strains isolated from a human fecal sample were Th17-inducing in at least two different mammalian hosts (mice and rats, Figures 7E and 7F); therefore, they may function in humans and may act as pathobionts, because they were derived from UC patients. However, we did not observe any inflammatory changes in the intestines of mice inoculated with the 20 strains (data not shown). Furthermore, mining of publicly available data from the MetaHIT project (Qin et al., 2010) revealed that only three out of the 20 strains were significantly increased in the microbiomes of UC and Crohn's disease subjects (Figure S7B). Moreover, the 20 strains were not abundant in mice inoculated with a fecal sample from subject UC4-2, in which Th17 cells were induced similarly to UC5-1 microbiota-associated mice (Figure S7C). Therefore, the 20 strains may not be essential components for Th17 induction nor contribute to the pathogenesis of IBD in humans. Our study design has some limitations in this context, including the limited number of subjects tested and the fact that robust activity in mice may not be physiologically relevant in humans and vice versa. Nevertheless, this is a clear demonstration of Th17 cell induction by human-associated bacterial species and therefore represents a significant step forward in dissecting human intestinal microbiota biology. The intestinal Th17 responses to commensal and pathogenic bacteria with shared EC-adhering characteristics have obvious clinical implications in the context of IBD, vaccine design, and probiotics.

EXPERIMENTAL PROCEDURES

Materials and experimental procedures are detailed in the Supplemental Experimental Procedures.

Animals and Microbes

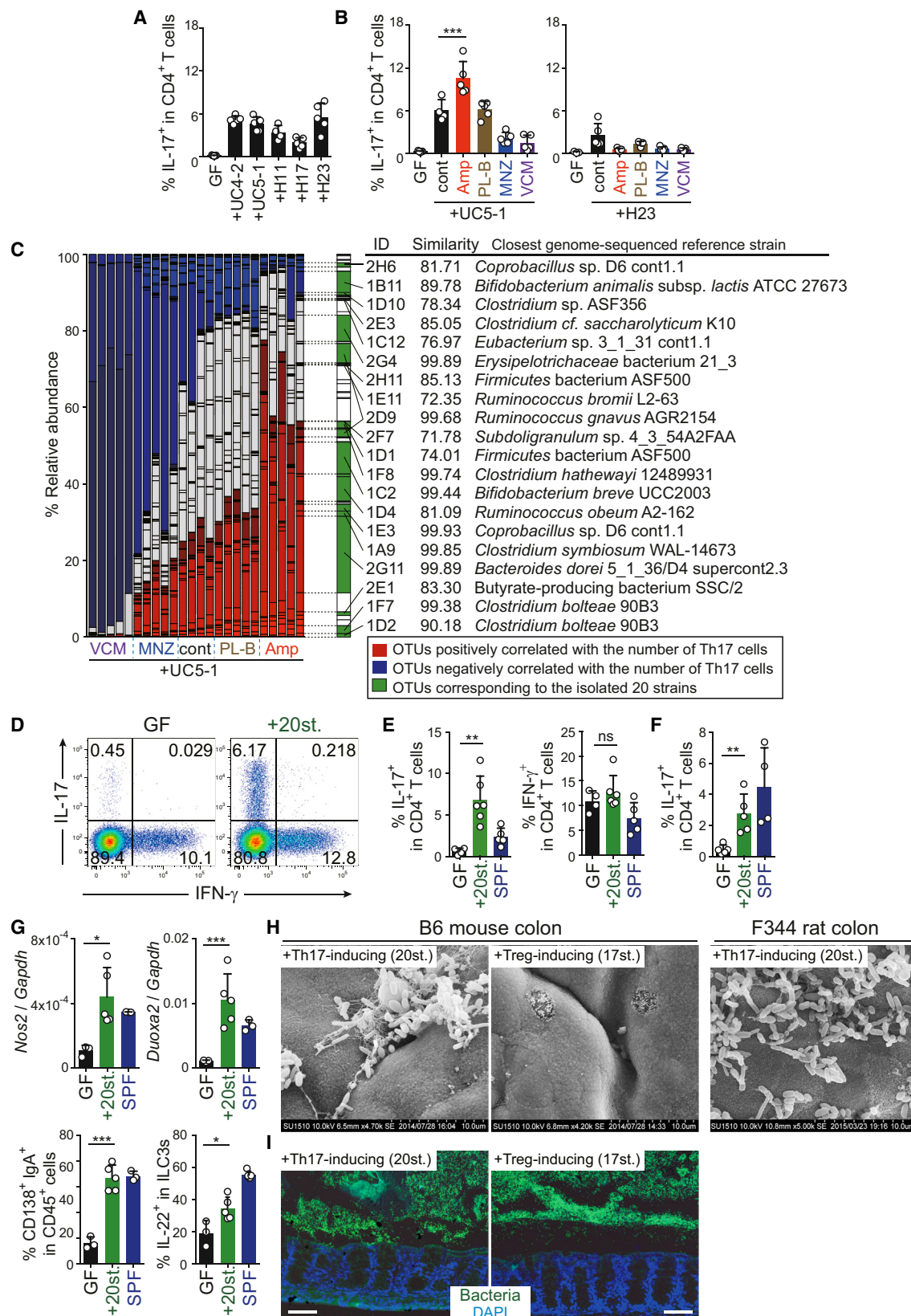
GF mice and rats were purchased from CLEA and Sankyo Labo Service and kept in the GF facility of Yakult Central Institute or RIKEN Yokohama Institute. To generate SFB-monocolonized animals, cecal or fecal suspensions from SFB-monocolonized animals were orally inoculated into GF animals. The *C. rodentium* Δ eae and *E. coli* O157:H7 Δ stx1 Δ stx2 mutant strains have been described previously (Kamada et al., 2012; Nagano et al., 2003). To generate the *E. coli* O157:H7 Δ stx1 Δ stx2 Δ eae mutant strain, the eae gene was replaced with the tetracycline-resistance gene by homologous recombination. Approximately 1×10^8 colony-forming unit (CFU) of *C. rodentium*, *E. coli* O157, or *L. monocytogenes* was orally administered into GF mice. *C. albicans* (TUA6) was provided by Dr. Takashi Umeyama (National Institute of Infectious Diseases, Japan) and orally administered ($\sim 1 \times 10^6$ CFU) into GF mice.

qPCR and 16S rRNA Gene Pyrosequencing

To quantify bacterial load, bacterial DNA was isolated from intestinal contents or mucosal tissues using a QIAamp Stool Mini Kit, and qPCR was carried out using universal or bacterial strain-specific primers for 16S rRNA genes. For 16S rRNA gene pyrosequencing, cecal contents were incubated with lysozyme and achromopeptidase. Sodium dodecyl sulfate and subsequently proteinase K was added to the suspension and high-molecular-mass DNA was purified by phenol/chloroform extraction. PCR was performed using primers to the V1-V2 region of the 16S rRNA gene, and the amplified DNA was used as template for 454 GS Junior sequencing.

Isolation of Human-Derived Th17-Inducing Bacterial Strains

Stool samples from patients with UC and from healthy donors were obtained and inoculated into IqI GF mice. The cecal contents from Amp-treated UC5-1



(legend on next page)

mice were cultured on Schaedler, BHI, GAM, CM0151, BL, or TS agar plates under strictly anaerobic conditions. Individual colonies were picked and identified by sequencing of the 16S rRNA gene fragment.

Statistical Analysis

All statistical analyses were performed using Prism software with two-tailed unpaired Student's *t* test or one-way ANOVA followed by Tukey's post hoc test. *p* values < 0.05 were considered significant (**p* < 0.05, ***p* < 0.01, and ****p* < 0.001).

ACCESSION NUMBERS

The accession number for the RNA sequencing data reported in this paper is NCBI GEO: GSE71734. The draft genome sequences of Th17-inducing 20 strains can be retrieved under DDBJ BioProject ID: PRJDB 4119-4138 and the 16S rRNA gene sequencing data under DDBJ BioProject ID: PRJDB4113.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.058>.

AUTHOR CONTRIBUTIONS

K.H. and Y.U. planned experiments, analyzed data, and wrote the manuscript. K.A., N.K., and T.T. performed immunological analyses and bacterial cultures together with Y.Nagano, S.N., Y.Noguchi, E.W., and H.Morita. W.S. and M.H. performed bacterial sequence analyses. M.A., S.N., and S.K. performed SEM analysis. A.I., H.S., T.N., E.I., T.Shima, and T.H. prepared M-SFB- and R-SFB-monocolonized animals and performed initial experiments. S.Y., S.T., H.Mori., and T.Sugiyama generated mutant strains of EHEC O157. I.I. and G.N. provided essential materials and contributed to data discussions. T.J. and H.O. provided *Il22r1^{fl/fl}* mice. T.K. helped with ChIP-seq experiments. K.T. instructed SEM. J.G.C. analyzed Saa1 regulatory regions.

ACKNOWLEDGMENTS

This work was supported by the Japan Society for the Promotion of Science NEXT program, Health Labor Sciences Research Grant, the Waksman Foundation of Japan, Suzuken Memorial Foundation, and the Uehara Memorial Foundation. We thank the animal care staff at the Yakult Central Institute, Osamu Ohara, Mayuko Sato, and Kei Hashimoto at RIKEN for their technical supports, Hiroyuki Nobusue and Hideyuki Saya at Keio University for data discussion, and Mamoru Totsuka at the University of Tokyo for providing the aMoS7 cell line.

Received: April 28, 2015

Revised: August 4, 2015

Accepted: August 17, 2015

Published: September 24, 2015

REFERENCES

- Atarashi, K., Nishimura, J., Shima, T., Umesaki, Y., Yamamoto, M., Onoue, M., Yagita, H., Ishii, N., Evans, R., Honda, K., and Takeda, K. (2008). ATP drives lamina propria T(H)17 cell differentiation. *Nature* 455, 808–812.
- Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K., et al. (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* 500, 232–236.
- Camp, J.G., Frank, C.L., Lickwar, C.R., Guturu, H., Rube, T., Wenger, A.M., Chen, J., Bejerano, G., Crawford, G.E., and Rawls, J.F. (2014). Microbiota modulate transcription in the intestinal epithelium without remodeling the accessible chromatin landscape. *Genome Res.* 24, 1504–1516.
- Chung, H., Pamp, S.J., Hill, J.A., Surana, N.K., Edelman, S.M., Troy, E.B., Reading, N.C., Villablanca, E.J., Wang, S., Mora, J.R., et al. (2012). Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* 149, 1578–1593.
- Conti, H.R., and Gaffen, S.L. (2010). Host responses to *Candida albicans*: Th17 cells and mucosal candidiasis. *Microbes Infect.* 12, 518–527.
- Croxen, M.A., and Finlay, B.B. (2010). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat. Rev. Microbiol.* 8, 26–38.
- Davis, C.P., and Savage, D.C. (1974). Habitat, succession, attachment, and morphology of segmented, filamentous microbes indigenous to the murine gastrointestinal tract. *Infect. Immun.* 10, 948–956.
- Derebe, M.G., Zlatkov, C.M., Gattu, S., Ruhn, K.A., Vaishnava, S., Diehl, G.E., MacMillan, J.B., Williams, N.S., and Hooper, L.V. (2014). Serum amyloid A is a retinol binding protein that transports retinol during bacterial infection. *eLife* 3, e03206.
- Dupont, S., Morsut, L., Aragona, M., Enzo, E., Giulitti, S., Cordenonsi, M., Zancato, F., Le Digabel, J., Forcato, M., Bicciato, S., et al. (2011). Role of YAP/TAZ in mechanotransduction. *Nature* 474, 179–183.
- Gaboriau-Routhiau, V., Rakotobe, S., Lécuyer, E., Mulder, I., Lan, A., Bridonneau, C., Rochet, V., Pisi, A., De Paepe, M., Brandi, G., et al. (2009). The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* 31, 677–689.
- Goto, Y., Panea, C., Nakato, G., Cebula, A., Lee, C., Diez, M.G., Laufer, T.M., Ignatowicz, L., and Ivanov, I.I. (2014). Segmented filamentous bacteria antigens presented by intestinal dendritic cells drive mucosal Th17 cell differentiation. *Immunity* 40, 594–607.
- Ivanov, I.I., Frutos, Rde.L., Manel, N., Yoshinaga, K., Rifkin, D.B., Sartor, R.B., Finlay, B.B., and Littman, D.R. (2008). Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* 4, 337–349.
- Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., et al. (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* 139, 485–498.
- Jepson, M.A., Clark, M.A., Simmons, N.L., and Hirst, B.H. (1993). Actin accumulation at sites of attachment of indigenous apathogenic segmented filamentous bacteria to mouse ileal epithelial cells. *Infect. Immun.* 61, 4001–4004.

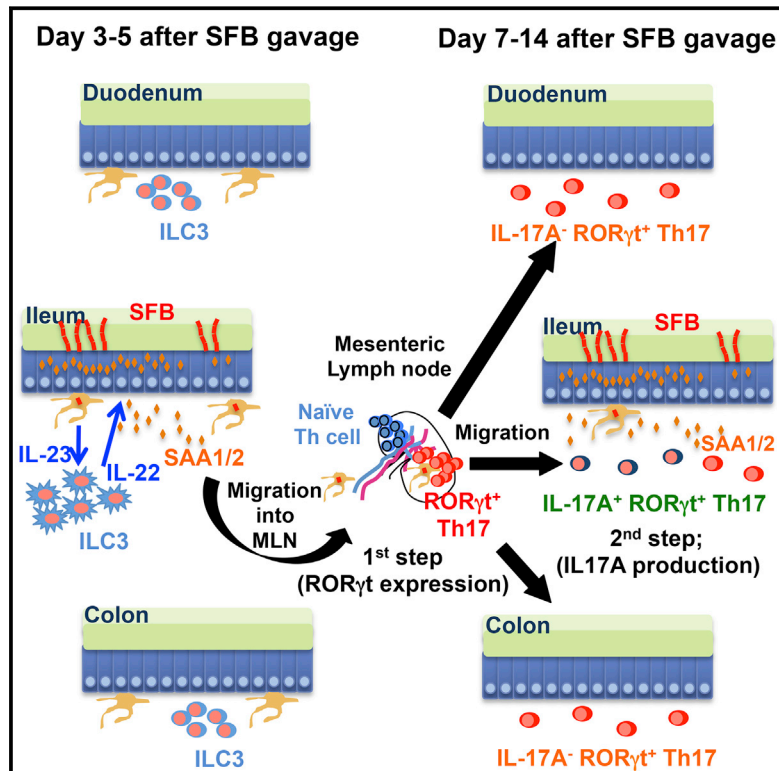
Figure 7. Th17 Induction and EC Adhesion by 20 Bacterial Strains Derived from the Human Intestine

(A) Th17 cell frequencies in IQI GF mice colonized with stool from patients with UC or healthy adults.
(B) Th17 cell frequencies in UC5-1 or H23-colonized IQI mice either untreated (cont) or treated with the indicated antibiotics.
(C) Cecal microbiota compositions of each mouse (*n* = 4 or 5 per group). OTUs positively correlated with the frequency of Th17 cells are marked in red and those negatively correlated marked in blue. OTUs corresponding to the isolated 20 strains are marked in green. ID, isolated strain ID. See also Table S1.
(D–F) The percentage of Th17 and Th1 cells in the colonic LP of B6 mice (D and E) and F344 rats (F) colonized with the mixture of 20 strains (20st.).
(G) qPCR for *Nos2* and *Duoxa2* in colonic EC (upper panels) and FACS for LP IgA⁺ cells and IL-22⁺ ILC3s (lower panels) of B6 mice colonized with the 20 strains.
(H and I) SEM images and FISH staining with EUB338 of the proximal colon of mice and rats colonized with the 20 Th17-inducing strains versus the 17 Treg-inducing strains. Scale bar, 100 μ m.
Error bars represent SD. See also Figure S7.

- Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. USA* 105, 15064–15069.
- Kamada, N., Kim, Y.G., Sham, H.P., Vallance, B.A., Puente, J.L., Martens, E.C., and Núñez, G. (2012). Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* 336, 1325–1329.
- Klaasen, H.L., Koopman, J.P., Van den Brink, M.E., Bakker, M.H., Poelma, F.G., and Beynen, A.C. (1993). Intestinal, segmented, filamentous bacteria in a wide range of vertebrate species. *Lab. Anim.* 27, 141–150.
- Mangan, P.R., Harrington, L.E., O'Quinn, D.B., Helms, W.S., Bullard, D.C., Elson, C.O., Hatton, R.D., Wahl, S.M., Schoeb, T.R., and Weaver, C.T. (2006). Transforming growth factor- β induces development of the T(H)17 lineage. *Nature* 441, 231–234.
- Mortha, A., Chudnovskiy, A., Hashimoto, D., Bogunovic, M., Spencer, S.P., Belkaid, Y., and Merad, M. (2014). Microbiota-dependent crosstalk between macrophages and ILC3 promotes intestinal homeostasis. *Science* 343, 1249288.
- Nagano, K., Taguchi, K., Hara, T., Yokoyama, S., Kawada, K., and Mori, H. (2003). Adhesion and colonization of enterohemorrhagic *Escherichia coli* O157:H7 in cecum of mice. *Microbiol. Immunol.* 47, 125–132.
- Palm, N.W., de Zoete, M.R., Cullen, T.W., Barry, N.A., Stefanowski, J., Hao, L., Degnan, P.H., Hu, J., Peter, I., Zhang, W., et al. (2014). Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* 158, 1000–1010.
- Pamp, S.J., Harrington, E.D., Quake, S.R., Relman, D.A., and Blainey, P.C. (2012). Single-cell sequencing provides clues about the host interactions of segmented filamentous bacteria (SFB). *Genome Res.* 22, 1107–1119.
- Prakash, T., Oshima, K., Morita, H., Fukuda, S., Imaoka, A., Kumar, N., Sharma, V.K., Kim, S.W., Takahashi, M., Saitou, N., et al. (2011). Complete genome sequences of rat and mouse segmented filamentous bacteria, a potent inducer of th17 cell differentiation. *Cell Host Microbe* 10, 273–284.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al.; MetaHIT Consortium (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65.
- Sano, T., Huang, W., Hall, J.A., Yang, Y., Chen, A., Gavzy, S.J., Lee, J., Ziel, J., Miraldi, E.R., Domingos, A.I., et al. (2015). An 23R/IL-22 circuit regulates epithelial serum amyloid A to promote local effector Th17 responses. *Cell* 163. Published online September 24, 2015. <http://dx.doi.org/10.1016/j.cell.2015.08.061>.
- Shaw, M.H., Kamada, N., Kim, Y.G., and Núñez, G. (2012). Microbiota-induced IL-1 β , but not IL-6, is critical for the development of steady-state TH17 cells in the intestine. *J. Exp. Med.* 209, 251–258.
- Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenko, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120.
- Spits, H., Artis, D., Colonna, M., Diefenbach, A., Di Santo, J.P., Eberl, G., Koyasu, S., Locksley, R.M., McKenzie, A.N., Mebius, R.E., et al. (2013). Innate lymphoid cells—a proposal for uniform nomenclature. *Nat. Rev. Immunol.* 13, 145–149.
- Stanley, A., Thompson, K., Hynes, A., Brakebusch, C., and Quondamatteo, F. (2014). NADPH oxidase complex-derived reactive oxygen species, the actin cytoskeleton, and Rho GTPases in cell migration. *Antioxid. Redox Signal.* 20, 2026–2042.
- Talham, G.L., Jiang, H.Q., Bos, N.A., and Cebra, J.J. (1999). Segmented filamentous bacteria are potent stimuli of a physiologically normal state of the murine gut mucosal immune system. *Infect. Immun.* 67, 1992–2000.
- Torchinsky, M.B., Garaude, J., Martin, A.P., and Blander, J.M. (2009). Innate immune recognition of infected apoptotic cells directs T(H)17 cell differentiation. *Nature* 458, 78–82.
- Umesaki, Y., Setoyama, H., Matsumoto, S., Imaoka, A., and Itoh, K. (1999). Differential roles of segmented filamentous bacteria and clostridia in development of the intestinal immune system. *Infect. Immun.* 67, 3504–3511.
- Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* 334, 255–258.
- Weaver, C.T., Elson, C.O., Fouser, L.A., and Kolls, J.K. (2013). The Th17 pathway and inflammatory diseases of the intestines, lungs, and skin. *Annu. Rev. Pathol.* 8, 477–512.
- Wu, H.J., Ivanov, I.I., Darce, J., Hattori, K., Shima, T., Umesaki, Y., Littman, D.R., Benoist, C., and Mathis, D. (2010). Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* 32, 815–827.
- Yang, Y., Torchinsky, M.B., Gobert, M., Xiong, H., Xu, M., Linehan, J.L., Alonzo, F., Ng, C., Chen, A., Lin, X., et al. (2014). Focused specificity of intestinal TH17 cells towards commensal bacterial antigens. *Nature* 510, 152–156.
- Yokoyama, S., Suzuki, T., Shiraishi, S., Ohishi, N., Yagi, K., Ichihara, S., Itoh, S., and Mori, H. (2001). Construction of deletion mutants of Shiga (-Like) toxin genes (stx-1 and/or stx-2) on enterohemorrhagic *Escherichia coli* (O157: H7). *J. Clin. Biochem. Nutr.* 30, 33–42.

An IL-23R/IL-22 Circuit Regulates Epithelial Serum Amyloid A to Promote Local Effector Th17 Responses

Graphical Abstract



Authors

Teruyuki Sano, Wendy Huang, Jason A. Hall, ..., Ana I. Domingos, Richard Bonneau, Dan R. Littman

Correspondence

dan.littman@med.nyu.edu

In Brief

Cytokines produced by Th17 cells are critical for protection and repair of mucosal barriers. IL-17A production is orchestrated by localization of specific poised Th17 cells at sites where the gut microbiome drives epithelial secretion of serum amyloid A.

Highlights

- SFB-specific Th17 cells are primed in MLN and traffic throughout the GI tract
- IL-17A-producing Th17 cells are enriched in ileum, where SFB adheres to epithelium
- SAA1/2 from ileal epithelial cells contributes the induction of IL-17A in Th17 cells
- SFB-induced IL-22 production by ILC3 stimulates epithelial cells to make SAA1/2

Accession Numbers

GSE70599

An IL-23R/IL-22 Circuit Regulates Epithelial Serum Amyloid A to Promote Local Effector Th17 Responses

Teruyuki Sano,^{1,7} Wendy Huang,^{1,7} Jason A. Hall,¹ Yi Yang,^{1,8} Alessandra Chen,^{1,2} Samuel J. Gavzy,¹ June-Yong Lee,¹ Joshua W. Ziel,¹ Emily R. Miraldi,^{1,3,4,5} Ana I. Domingos,⁶ Richard Bonneau,^{3,4,5} and Dan R. Littman^{1,2,*}

¹Molecular Pathogenesis Program, The Kimmel Center for Biology and Medicine of the Skirball Institute, New York University School of Medicine, New York, NY 10016, USA

²The Howard Hughes Medical Institute, New York University School of Medicine, New York, NY 10016, USA

³Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY 10003, USA

⁴Courant Institute of Mathematical Sciences, Computer Science Department, New York University, New York, NY 10003, USA

⁵Simons Center for Data Analysis, Simons Foundation, New York, NY 10010, USA

⁶Obesity Laboratory, Instituto Gulbenkian de Ciência, Oeiras 2780-156, Portugal

⁷Co-first author

⁸Present address: Department of Microbiology and Immunology, Medical University of South Carolina, 173 Ashley Avenue, Charleston, SC 29425, USA

*Correspondence: dan.littman@med.nyu.edu

<http://dx.doi.org/10.1016/j.cell.2015.08.061>

SUMMARY

ROR γ t⁺ Th17 cells are important for mucosal defenses but also contribute to autoimmune disease. They accumulate in the intestine in response to microbiota and produce IL-17 cytokines. Segmented filamentous bacteria (SFB) are Th17-inducing commensals that potentiate autoimmunity in mice. ROR γ t⁺ T cells were induced in mesenteric lymph nodes early after SFB colonization and distributed across different segments of the gastrointestinal tract. However, robust IL-17A production was restricted to the ileum, where SFB makes direct contact with the epithelium and induces serum amyloid A proteins 1 and 2 (SAA1/2), which promote local IL-17A expression in ROR γ t⁺ T cells. We identified an SFB-dependent role of type 3 innate lymphoid cells (ILC3), which secreted IL-22 that induced epithelial SAA production in a Stat3-dependent manner. This highlights the critical role of tissue microenvironment in activating effector functions of committed Th17 cells, which may have important implications for how these cells contribute to inflammatory disease.

INTRODUCTION

The vertebrate gastrointestinal (GI) tract is colonized by hundreds of distinct species of microorganisms that collectively maintain a mutualistic relationship with the host (Macpherson and Harris, 2004). This mutualism is critically dependent on a state of balanced immune activation, which fosters cohabitation between the host and microbiota, while providing optimal protection against opportunistic pathogens (Honda and Littman, 2012). It is now appreciated that the composition of the micro-

biome can contribute significantly to this immunological balance, in part through the capacity of individual bacterial or viral species to profoundly influence distinct arms of the immune response by themselves or in concert with other microbial species (Hooper et al., 2012; Virgin, 2014). For instance, *Bacteroides fragilis* and mixtures of various strains of *Clostridia*, which colonize the large intestine, have been shown to promote intestinal and systemic immune tolerance through regulatory T cells (Atarashi et al., 2013; Lathrop et al., 2011; Round et al., 2011), and segmented filamentous bacteria (SFB) induce antigen-specific T-helper-17 cells (Th17) cell differentiation in the small intestine (Gaboriau-Routhiau et al., 2009; Ivanov et al., 2009; Yang et al., 2014).

Th17 cells constitute a subset of activated CD4⁺ T cells that are characterized by the production of the cytokines interleukin (IL)-17A, IL-17F, and IL-22 (Korn et al., 2009). These cytokines act on a broad range of hematopoietic and non-hematopoietic cells to regulate central aspects of host immunity, including granulopoiesis, neutrophil recruitment, and the induction of antimicrobial peptides (Weaver et al., 2007). Th17 cells thus execute important functions in both epithelial homeostasis and host defense against various extracellular pathogens, such as *Candida albicans* and *Pseudomonas aeruginosa* (Stockinger and Veldhoen, 2007). Conversely, over-exuberant Th17 responses may promote auto-inflammatory diseases, such as Crohn's disease, rheumatoid arthritis (RA), psoriasis, and multiple sclerosis (MS) (Furuzawa-Carballeda et al., 2007). While genetic polymorphisms significantly factor into the onset of these diseases, emerging evidence also highlights the influence that environmental factors, such as diet and microbial composition, can impose on such propensities. Accordingly, recent studies have illustrated the potential of SFB to exacerbate Th17-mediated disease in murine autoimmune models of both RA and MS (Lee et al., 2011; Wu et al., 2010), although the intermediate molecular steps connecting SFB to a distal immune response are ill defined.

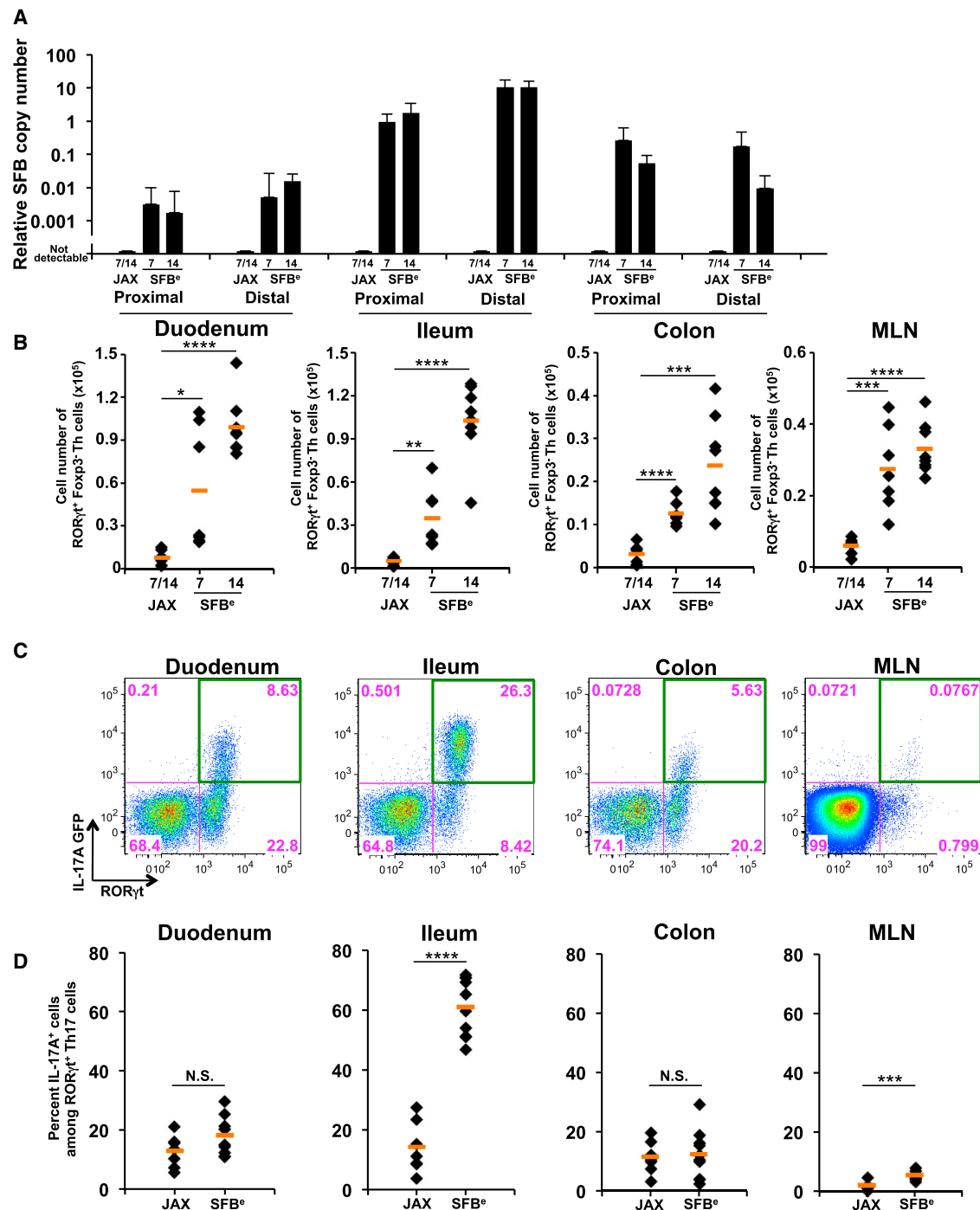


Figure 1. Site of SFB Colonization Correlates with Maximal IL-17A Induction in RORγt⁺ T Cells

(A) Quantitative PCR (qPCR) analysis of SFB genomic 16S in duodenum, ileum, and colon (proximal and distal segments of each) from SFB-colonized (SFB^o; gavaged with SFB-rich fecal contents) or SFB-free (JAX; gavaged with SFB-free fecal contents) mice at 7 and 14 days post-gavage. Copy number of SFB 16S was normalized by copy number of mouse genomic *Il-23r* gene. The experiment was repeated more than three times with similar results combined. Data are represented as mean with SD.

(B) Number of RORγt⁺ CD4⁺ T cells in intestinal lamina propria (LP) or MLN of SFB-colonized (SFB^o; n = 7) or SFB-free (JAX; n = 7) mice on 7 and 14 days post-gavage. CD4⁺ T cells were isolated from the indicated tissues, and Foxp3⁻ RORγt⁺ CD4⁺ T cells were quantified by fluorescence-activated cell sorting (FACS).

(legend continued on next page)

SFB colonization of the small intestine promoted global transcriptional changes in host epithelia, including the induction of antimicrobial peptides and stress response genes, such as serum amyloid A (SAA1 and SAA2) (Ivanov et al., 2009). SAA is typically induced in response to infection and acute injury and can promote inflammation, in part through elicitation of proinflammatory cytokine production and recruitment of granulocytes, monocytes, and T lymphocytes (Uhlir and Whitehead, 1999). The effect of SAA on the immune response is context driven (Cray et al., 2009; Eckhardt et al., 2010; Ivanov et al., 2009), much like that of Th17 cells. Insofar as SFB and Th17 cells are intertwined, the question of whether SAA impacts aspects of Th17 biology remains to be addressed.

Th17 cells along with several other innate-like cell lineages, including specific subsets of $\gamma\delta$ T cells ($\gamma\delta 17$) and type 3 innate lymphoid cells (ILC3), are regulated by the transcription factor RAR-related orphan receptor gamma (ROR γ t) (Chien et al., 2013; Ivanov et al., 2006; Spits and Di Santo, 2011). However, in contrast to the requirement for antigen recognition in the context of major histocompatibility complex (MHC) to drive Th17 cell activation, $\gamma\delta 17$ and ILC3 effector functions are elicited independently of antigen presentation. For example, the proinflammatory cytokine IL-23 triggers rapid IL-17 and IL-22 secretion by $\gamma\delta 17$ cells and ILC3s, respectively, upon ligation of the highly expressed IL-23 receptor (IL-23R). As $\gamma\delta 17$ cells and ILC3s often reside in proximity to exposed mucosal surfaces, their activation typically precedes antigen-specific T cell differentiation and recruitment (Martin et al., 2009; Sutton et al., 2009). Whether this has bearing on the function of newly recruited T cells is unclear.

We utilized models of acute SFB colonization to investigate the mechanism of Th17 cell induction in a spatiotemporal context. We found that, following early induction of SFB-specific ROR γ t⁺ Th17 cells in the mesenteric lymph nodes, there was distribution of such cells throughout the length of the gut, from duodenum to colon, but IL-17A expression was largely confined to the terminal ileum, the site of SFB attachment to epithelium (Ivanov et al., 2009). We have identified an SFB-triggered circuit in which ILC3 secretion of IL-22 is critical for local epithelial production of SAA1 and SAA2, which act directly on poised Th17 cells to amplify effector cytokine production. These findings suggest that tissue microenvironments contribute to the acquisition of effector functions by polarized activated effector and memory cells.

RESULTS

Selective IL-17A induction in ROR γ t⁺ T Cells in Ileum of SFB-Colonized Mice

SFB colonization results in a striking increase in both the number and proportion of Th17 cells among total CD4⁺ T cells within the

small intestine lamina propria (SILP) (Ivanov et al., 2009). To explore the mechanism of Th17 cell induction, we orally introduced fecal contents from *Il-23^{gfp/gfp} Rag2^{-/-}* mice, which have about 100-fold enriched fecal SFB (SFB^e) compared to wild-type (WT) animals (Yang et al., 2014), into SFB-free (JAX) mice and subsequently traced the kinetics of Th17 cell differentiation in different compartments of the GI tract on days 7 and 14. As expected, SFB colonization was most abundant in the terminal ileum (Figures 1A and S1A). After 5 days of SFB^e gavage, we observed a marked increase in proportion and number of CD4⁺ T cells expressing ROR γ t, which defines the Th17 cell program, in the mesenteric lymph nodes (MLN) that drain the intestine (Figure S1B). Adoptively transferred carboxyfluorescein succinimidyl ester (CFSE)-labeled SFB-specific T cells from 7B8 T cell receptor transgenic mice (Yang et al., 2014) proliferated and upregulated ROR γ t expression in the MLN by day 4 after SFB colonization (Figure S1C). Expansion of these cells was then observed throughout the SILP by 7 or 14 days post-SFB gavage (Figures 1B and S1C). Using fluorescently labeled MHC class II tetramers with a bound immunodominant SFB antigen, we found SFB-specific ROR γ t⁺ cells throughout the gut; however, they were enriched in the ileum, possibly reflecting additional expansion in response to local antigen (Figure S1D).

To track expression of IL-17A, we colonized *Il17a^{gfp/gfp}* reporter mice with SFB^e. Strikingly, co-staining of ROR γ t and GFP indicated that the majority of ROR γ t⁺ cells expressed IL-17A mRNA only in the terminal ileum, but not in other regions of the GI tract (Figures 1C and 1D). Induction of IL-17A in poised Th17 cells thus appears to occur primarily at the site of SFB colonization.

Serum Amyloid A Proteins Modulate Th17 Cell Cytokine Production

Recent studies have shown that the tissue microenvironment confers distinct transcriptional and functional properties to otherwise similar lymphoid and myeloid cells (Burzyn et al., 2013; Cipolletta et al., 2012; Lavin et al., 2014; Lewis et al., 2011). Our observation of the spatial restriction of IL-17A-producing Th17 cells raised the possibility that SFB induces changes in the microenvironment of the underlying lamina propria, thus contributing to the differentiation state of poised Th17 cells. Consistent with this possibility, germ-free (GF) mice monocolonized with SFB upregulated expression of several secreted molecules in the terminal ileum, among which SAA1 and SAA2 (referred to as SAA1/2, as they are highly homologous) were the most strongly induced (Ivanov et al., 2009). Because SFB attaches to the epithelium, we focused on secreted molecules that changed expression within ileal epithelial cells for their potential influence on the functional status of adjacent Th17 cells. To identify epithelium-specific translated RNAs, we employed in vivo translating ribosome affinity purification

(C) IL-17A (GFP⁺) production in Foxp3⁻ ROR γ t⁺ CD4⁺ T cells from intestinal LPs and MLN in IL-17A reporter (*Il17a^{gfp/gfp}*) mice at 14 days post-SFB colonization. IL-17A-producing Th17 cells were monitored by FACS after staining with anti-GFP antibody without re-stimulation.

(D) The proportion of IL-17A producing Th17 cells among Foxp3⁻ ROR γ t⁺ CD4⁺ T cells following SFB colonization. Each dot represents a single mouse, and the orange bar indicates average of relative gene expression of each group. Data were from at least two independent experiments (JAX; n = 7. SFB^e; n = 9). N.S., not significant (p > 0.05); *p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001. See also Figure S1.

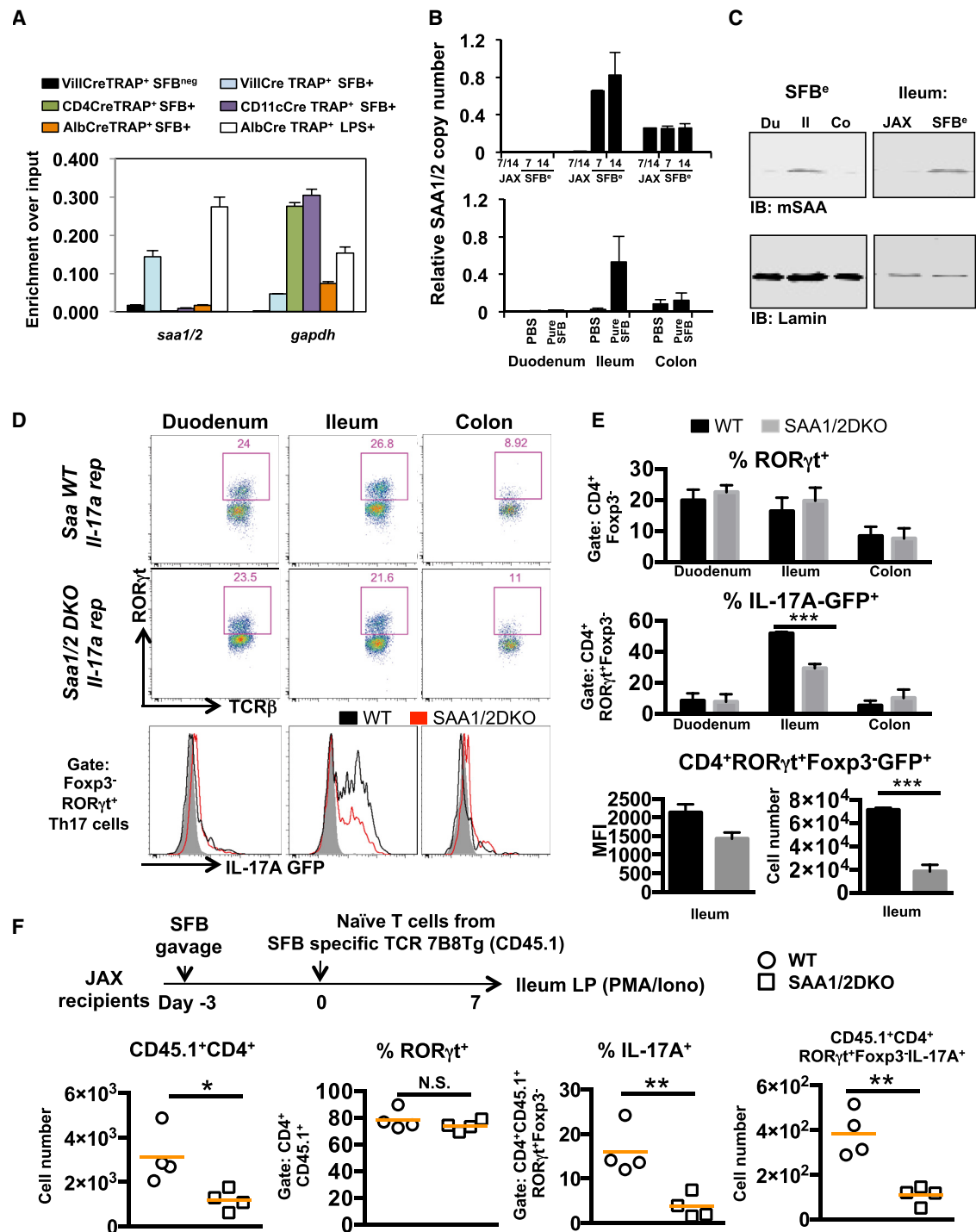


Figure 2. Role of Epithelial SAA1/2 in Ileal Th17 Cell Production of IL-17A

(A) Translating ribosome affinity purification (TRAP) assay for identification of the cell type expressing SAA1/2. Four individual lines of TRAP mice, conditionally expressing a GFP tagged ribosome L10 fusion protein in intestinal epithelial (Villin-Cre), T (CD4-Cre), myeloid (CD11c-Cre), or liver (Albumin-Cre) cells were generated. After 2 weeks of SFB colonization, ribosomes in small intestinal lysates collected from each line were enriched by GFP pull-down. Actively translating SAA1/2 mRNAs were quantified by qRT-PCR. Black and white bar indicates TRAP signal from SFB-free animals and light blue, dark blue, purple, and orange bars indicate TRAP signal from SFB-positive animals. White bar indicates TRAP signal from tissue harvested from JAX animals at 6 hr following intraperitoneal administration of LPS.

(B) Quantification of SAA1/2 mRNA in intestinal epithelial cells of duodenum, ileum, and colon from SFB⁺ (day 7; n = 8, day 14; n = 3), pure SFB (n = 4), JAX (n = 3), or PBS-gavaged (n = 8) mice. Data were normalized using Gapdh. Similar experiments were performed three times.

(legend continued on next page)

(TRAP), which was developed to identify cell type-specific transcripts in tissues containing heterogeneous cell populations (Heiman et al., 2014). TRAP Villin-Cre, TRAP CD11c-Cre, TRAP CD4-Cre, and TRAP Albumin-Cre mice expressing a conditionally induced GFP ribosome L10 fusion protein from the *rp10* locus in epithelial cells, dendritic cells, T cells, and hepatocytes, respectively, were colonized with SFB for 2 weeks or left untreated, after which GFP pull-down of tissue lysates and qRT-PCR using transcript-specific primers were performed to identify translating mRNAs in specific cell types. Using primers that amplify both SAA1 and SAA2 mRNAs, we found that both were highly transcribed and translated in epithelial cells from SFB⁺ mice (Figure 2A), consistent with previous reports of SAA protein expression in small and large intestinal epithelial cells (Derebe et al., 2014; Eckhardt et al., 2010). We next assessed the spatiotemporal induction of SAA1/2 mRNAs in the epithelial fraction of the intestine upon acute colonization with SFB (Figure 2B). Strikingly, SAA1/2 mRNA induction was confined to the terminal ileum and was modulated in synchrony with SFB colonization (Figures 1A and S1A). Differential SAA expression was confirmed by immunoblotting of the EDTA-epithelial cell-rich fractions from different regions of the GI tract of animals with or without SFB colonization (Figure 2C). Notably, this localization corresponds to the site of SFB attachment and the location of IL-17A-producing Th17 cells.

We next asked whether SAA1/2 contribute to functional differentiation of Th17 cells in vivo. Wild-type (WT) or SAA1 and SAA2 (DKO) mutant mice (Eckhardt et al., 2010) were crossed with *Il17a^{gfp/gfp}* reporter mice in an SFB-free background. Seven days after SFB⁺ gavage, WT and DKO littermates were colonized at similar levels (Figure S2A). On day 7, at the peak of the Th17 response, we observed in WT and DKO animals comparable proportions of Foxp3^{neg} RORγt⁺ CD4⁺ T cells in different segments of the GI tract. Within the terminal ileum, however, production of IL-17A-GFP in the poised Th17 cells was significantly reduced in DKO mice, both proportionally and on a per-cell basis (Figures 2D and 2E). In addition, IL-17A and IL-17F mRNAs in the SILP of DKO animals were significantly reduced relative to WT littermates (Figure S2A). Unlike IL-17A and IL-17F, the chemokine receptor CCR6, a marker of Th17 cells, was expressed in similar amounts in wild-type and SAA1/2DKO animals (Figure S2B), consistent with SAA1/2 contributing to a specific subset of genes in the Th17 cell program.

To exclude a possible role for SAA1/2 in T cell development or repertoire bias, we colonized WT and DKO hosts with SFB⁺ feces and adoptively transferred into them naive CD4⁺ T cells from 7B8 transgenic mice (Yang et al., 2014). One week post-transfer, we noted that comparable proportions of donor-derived cells

became RORγt⁺ within the terminal ileum of both DKO and WT animals. However, production of IL-17A by the RORγt⁺ 7B8 cells upon PMA/Ionomycin restimulation was markedly attenuated in DKO recipients (Figure 2F). Similar results were observed in WT and DKO recipients of T cells from 7B8 transgenic mice bred onto the IL17A-IRES-GFP reporter background (Figure S2C). Although almost all SFB-specific donor-derived T cells found in the ileum of SFB-colonized mice expressed RORγt regardless of the host genotype, we observed a reduction in total number of these cells recovered in SAA1/2 DKO animals (Figure 2F). This suggests that, in addition to a role in promoting Th17 cytokine production, SAA may promote local Th17 cell proliferation and/or retention.

Next, we treated murine naive CD4⁺ T cells cultured in Th17 polarizing conditions with recombinant mouse SAA1 (rmSAA1). SAA1 significantly enhanced IL-17A and IL-17F production in RORγt⁺ CD4⁺ T cells cultured in suboptimal Th17 polarizing conditions (Figures 3A–3C and S3A). Importantly, this effect was reversed when anti-mSAA blocking antibody was introduced (Figure S3A). In contrast, CCR6 expression was not changed by rmSAA1 treatment (Figure S3B). Based on our in vivo observation, we also asked whether mSAA1 exerts a direct effect on cell proliferation in addition to Th17 signature cytokine production. Labeling of naive CD4⁺ T cells with CFSE and intracellular Ki67 staining showed enhanced proliferation of RORγt⁺ Th17 cells upon treatment with rmSAA1 (Figures 3A–3C and S3C), but the higher proportions of cells expressing IL-17A and IL-17F were observed independently of the extent of proliferation (Figure 3C). Furthermore, in vitro differentiation of naive CD4⁺ T cells toward Th1, Th2, and iTregs was not significantly affected in the presence of recombinant mouse SAA1 (Figure S3D). Similarly, rhSAA augmented IL-17A production during human Th17 cell priming (Figure S3E). Importantly, this effect was reversed when SAA blocking antibody was introduced into culture. These results suggest that SAA acts directly on both mouse and human Th17 cells to enhance their differentiation and effector function.

To further study how SAA influences the transcriptional program of Th17 cells, we performed RNA sequencing (RNA-seq) with control and rmSAA1-treated murine Th17 cells. We identified a specific subset of the Th17 gene expression program that was specifically affected by rmSAA1 treatment (Figures 3D and S3F). Unbiased Ingenuity Pathway Analysis predicted RORγt, the key regulator of the Th17 differentiation program, as the top upstream regulatory factor sensitive to SAA perturbation in vitro (Figure 3E). ROR-element-driven luciferase assays in Th17 cells confirmed that RORγt transcriptional activity was potentiated during rmSAA1 stimulation in vitro (Figure 3F).

(C) Representative western blots of lysates from the EDTA-epithelial-rich fraction of different regions of the GI tract in animals with or without SFB colonization. (D–F) Role of SAAs in intestinal LP Th17 cell differentiation following SFB colonization. (D and E) IL-17A reporter littermates in WT (black, n = 5) or *Saa1/2* double-knockout (DKO) (gray, n = 5) backgrounds were colonized with SFB for 7 days. RORγt⁺ cells among CD4⁺ LP T cells and their IL-17A producing capacity were measured by FACS. MFI and cell numbers of CD4⁺RORγt⁺Foxp3^{neg}GFP⁺ Th17 cells were quantified. (F) Naive CD4⁺ T cells from SFB-specific TCR transgenic (7B8) mice were adoptively transferred into SFB-colonized WT (circles, n = 4) or *Saa1/2* DKO (squares, n = 4) littermates (as indicated in diagram). Total number of donor cells recovered from ileum (far left), percentage of RORγt-expressing donor-derived cells (center left), percentage of IL-17A-expressing cells upon ex vivo re-stimulation by PMA/Ionomycin (center right), and their total number (far right) were monitored by FACS at 7 days post-transfer. N.S., not significant (p > 0.05); *p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001. Data are represented as mean with SD. See also Figure S2.

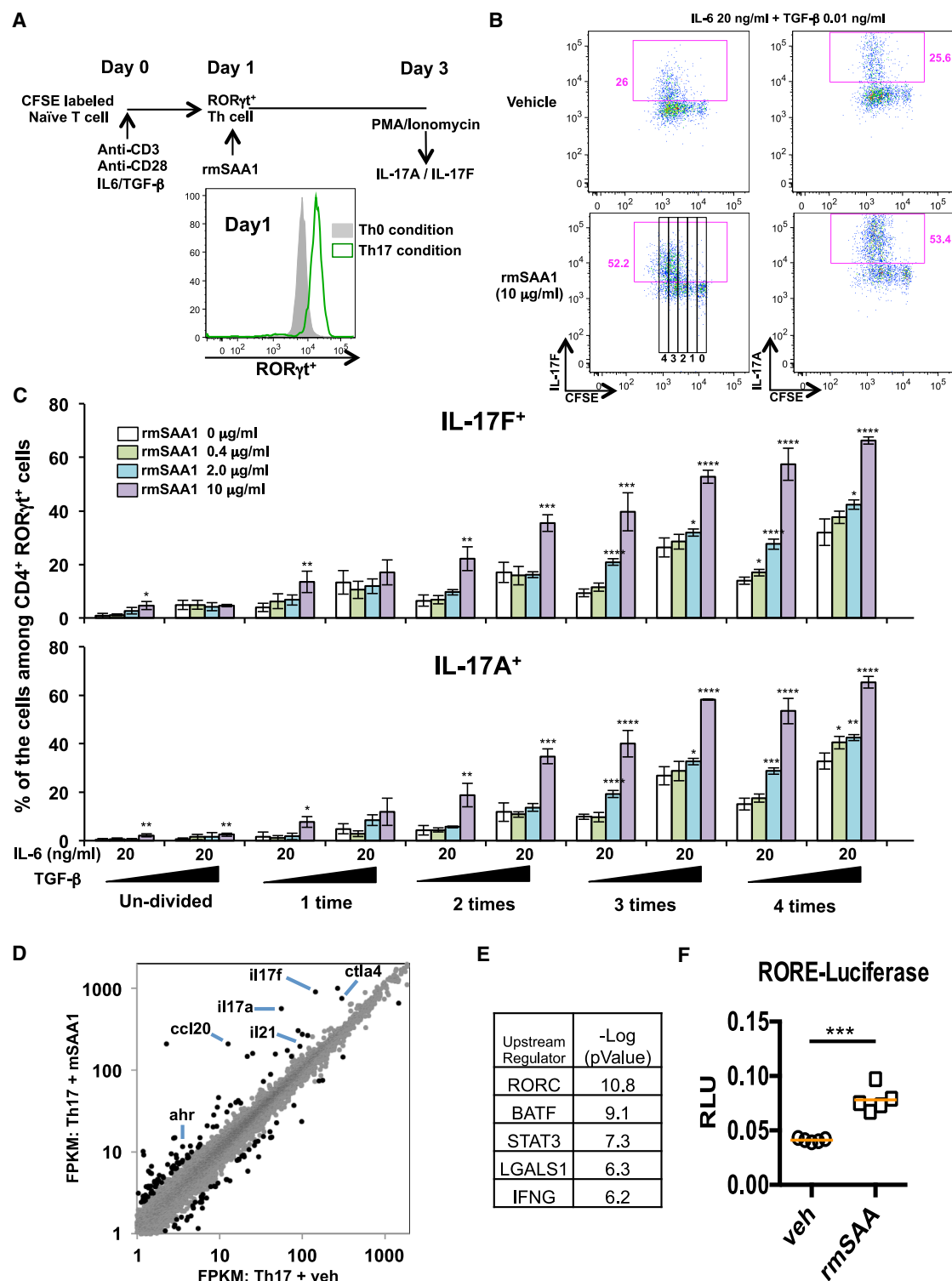


Figure 3. SAA1 Promotes IL-17A Production in Polarized Th17 Cells In Vitro

(A) Schematic representation of the experimental protocol. CFSE-labeled naive CD4⁺ T cells from WT mice were activated under non-polarizing (Th0: anti-CD3 and anti-CD28) or Th17 polarization condition (Th17: anti-CD3, anti-CD28, IL-6, and TGF-β). At 24 hr, RORγt⁺ T cells were further cultured with recombinant mouse SAA1 (rmSAA1) under Th17 polarization conditions for 48 hr. (B) IL-17A and IL-17F production by in vitro-polarized mouse Th17 cells (CD4⁺ RORγt⁺) at 48 hr after rmSAA1 (10 μg/ml) addition. Cells were re-stimulated with PMA/Ionomycin for 5 hr, and cytokine production and CFSE dilution were monitored by FACS analysis.

(legend continued on next page)

SFB-Induced SAA1/2 Requires IL-22 Expression and Epithelial Stat3 Signaling

We next sought to understand the upstream signaling pathway that regulates ileal SAA induction in response to SFB colonization. SFB-free C57BL/6 mice were co-housed with SFB-positive mice (Taconic) for 2 weeks before epithelial-specific translating RNAs were assessed. RNA-seq profiling of polysome-associated mRNAs pulled-down from SFB-colonized TRAP Villin-Cre mice revealed a significant enrichment of actively translated genes that are known to be regulated by the transcription factor Stat3 (Figure S4A). To corroborate this analysis, we used immunofluorescence to probe for Stat3 activity (pStat3) in small intestinal tissue sections of mice acutely colonized with SFB⁺ feces. We detected pStat3 within 24 hr of SFB colonization (Figure 4A). Strikingly, SFB⁺ fecal-induced upregulation of SAA1/2 was abrogated almost entirely in mice with intestinal epithelial cell-specific inactivation of Stat3 (Figure 4B). Stat3 activation is elicited by multiple cytokines, including IL-10 and members of the IL-20 subfamily, of which IL-22 is the best characterized (Rutz et al., 2014). Interestingly, we detected a significant increase in IL-22 mRNA within the terminal ileum shortly after colonization with SFB⁺, and this preceded the induction of epithelial SAA1/2 expression by one day (Figures 4C–4E). This result suggested that IL-22 activated Stat3, which then induced SAA1/2 expression. Indeed, treatment with a neutralizing anti-IL-22 antibody prevented SAA1/2 induction following SFB colonization (Figure S4B). Consistent with this result, SAA1/2 mRNA induction by SFB was almost completely abrogated in IL-22 deficient mice (Figure 4F). In addition, IL-17A production among ROR γ t⁺ Th17 cells in ileum was partially blocked in animals receiving IL-22 blocking antibodies, similar to what was observed in SAA1/2 DKO mice (Figure S4C).

To evaluate whether IL-22 can directly induce epithelial SAA1/2 expression, we generated small intestinal epithelial (SIE) organoids as previously described (Sato and Clevers, 2013) and exposed them to recombinant murine IL-22 (rIL-22). The addition of rIL-22 to SIE organoid cultures induced SAA1/2 mRNA in a dose-dependent fashion within 2 hr (Figure 4G). Coupled with the in vivo results, these findings indicate that SFB colonization initiates an IL-22-mediated Stat3 signaling cascade in epithelial cells, resulting in localized SAA1/2 expression.

SFB Stimulates IL-23R-Dependent IL-22 Production by Activated ILC3

In the intestine, IL-22 is produced by ROR γ t-expressing lymphoid cells that include ILC3, Th17, and $\gamma\delta$ 17 cells (Sawa et al., 2011). To investigate the source of IL-22 that directs

epithelial SAA production following SFB colonization, we compared RNA expression profiles from small intestine Th17 cells and ILC3 in mice stably colonized with SFB. Based on normalized DESeq reads, IL-22 expression was 10-fold higher in ILC3s than in Th17 cells (Figure S5A). Examination of these data sets also revealed a surface antigen, killer cell lectin-like receptor subfamily B member 1B (*Klrb1b*), which was uniquely expressed in ILC3. We confirmed this finding at the level of protein by flow cytometry (Figure S5B), and subsequently incorporated Klrb1b into the fluorescent cell-labeling panel to dispense with the necessity of a reporter in studying the dynamic behavior of ILC3 upon SFB⁺ colonization.

At day 4 of SFB⁺ colonization, up to 80% of ILC3s expressed the activation marker stem cell antigen-1 (Sca-1), which amounted to a 10-fold increase compared to the proportion of Sca-1⁺ ILC3s in mice gavaged with SFB-negative feces (Figures 5A and 5B). Unlike CCR6[−] ILC3s, CCR6⁺ ILC3s typically populate the isolated lymphoid follicles and cryptopatches (Lügering et al., 2010). Notably, both CCR6[−] and CCR6⁺ ILC3s expressed high level of Sca-1 on their surface in response to SFB colonization. Although ILC3s increased in size in response to SFB colonization, their absolute numbers remained unchanged (Figures S5C–S5E). Intracellular cytokine staining also revealed a striking increase in both the frequency and total number of ILC3s producing IL-22 (Figure 5C). We recapitulated these findings using the fecal contents from SFB mono-associated mice, supporting the notion that SFB alone is sufficient to direct the localized ILC3 response (Figure S5F). Importantly, at this early time point, other ROR γ t-dependent lineages, including $\gamma\delta$ 17 and (non-SFB-specific) Th17 cells, did not produce IL-22 (Figure 5D). Furthermore, epithelial cells from *Rag2* mutant mice, harboring ILC3s but lacking B and T cells, also expressed SAA1/2 in response to SFB colonization (Figure 5E).

ILC3s residing in the colon secrete IL-22 in response to IL-23 (Eken et al., 2014). To determine whether this mechanism also governs terminal ileal production of IL-22 by ILC3s upon SFB colonization, we colonized IL-23R-sufficient and -deficient mice with SFB⁺ feces and examined the spontaneous release of IL-22 by ILC3s. Both the frequency and absolute number of IL-22-secreting CCR6⁺ and CCR6[−] ILC3s declined sharply in IL-23R-deficient animals (Figures 6A and 6B). Consistent with this, induction of epithelial SAA1/2 mRNA was markedly blunted in IL-23R-deficient animals (Figure 6C). Altogether, these findings indicate that IL-23 signaling initiates IL-22-mediated cross-talk between ILC3s and small intestine epithelial cells, resulting in spatially restricted SAA production, which contributes to local augmentation of Th17 effector functions (Figure S6).

(C) Representative data from polarized cells (48 hr) (four experiments total). The frequencies of IL-17A- and IL-17F-producing cells during each cell division are displayed using the gates represented in (B). Data are shown as mean with SD.

(D) RNA-seq studies from two biological replicates comparing in vitro Th17 cells cultured in 20 ng/ml IL-6 and 0.1 ng/ml TGF- β with or without 2 μ g/ml rmSAA1 for 48 hr. Dots represent FPKMs of individual mRNA transcripts in the respective cell type (gray). 195 genes that were significantly different between vehicle and SAA treatment (p value <0.05, fold change >2) are highlighted in black.

(E) Ingenuity Pathway Analysis for upstream regulators of the 195 SAA-affected genes in Th17 cells.

(F) A ROR-responsive element (RORE) firefly luciferase reporter was used to monitor ROR γ t transcriptional activity in Th17 cells. The RORE-luciferase reporter construct was introduced by nucleofection into Th17 cells polarized for 48 hr and cultured with or without rmSAA1 (2 μ g/ml), and luciferase activity was measured 24 hr later. Data were normalized by renilla reporter activity. Results from four technical repeats in each condition were averaged.

Results are representative of three independent experiments. *p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001. See also Figure S3.

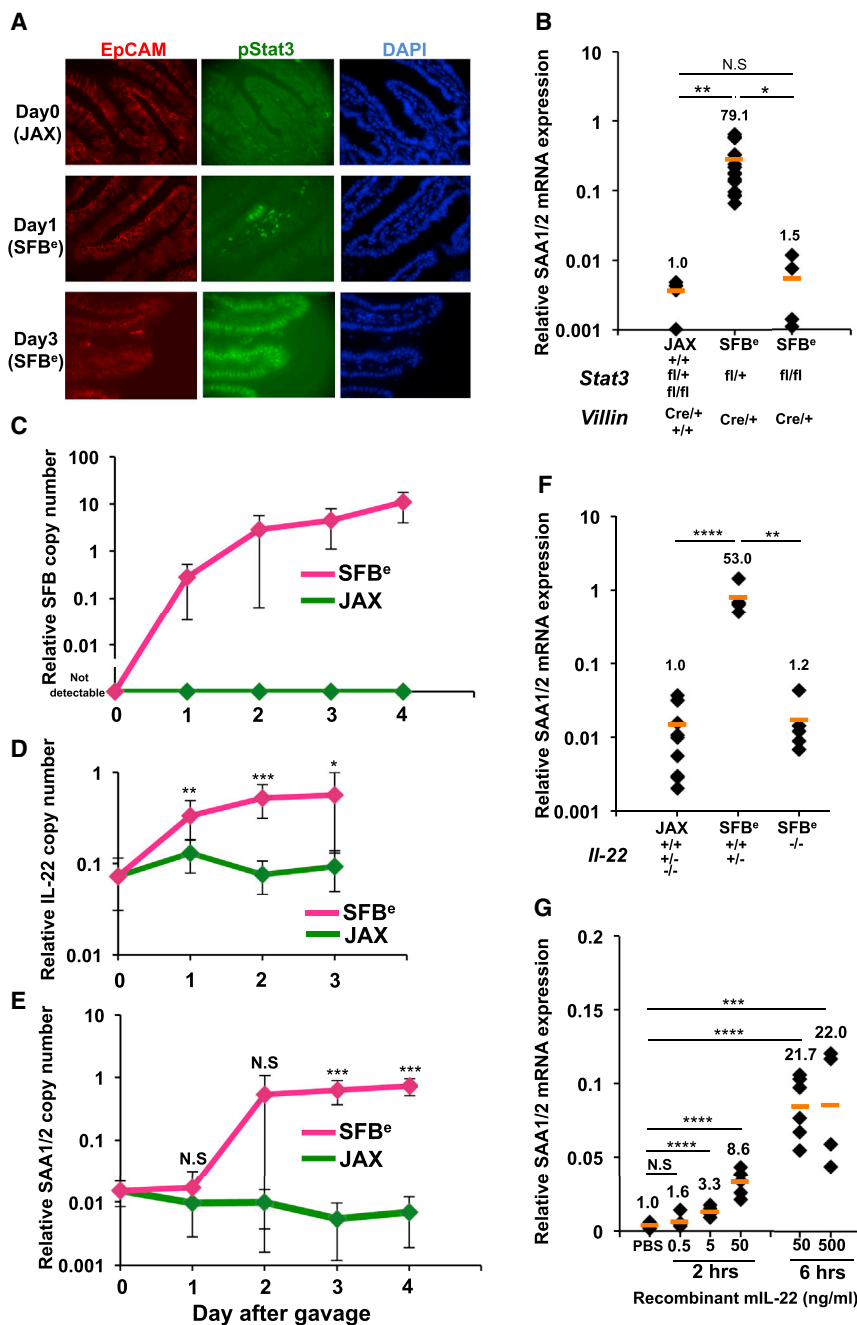


Figure 4. Requirement for IL-22 Activation of Epithelial Stat3 in SFB-Dependent Induction of SAA1/2

(A) Immunohistochemistry analysis of phospho-Stat3 (green) in terminal ileum from SFB-colonized (SFB^e) or SFB free (JAX) mice on day 1 or day 3 following gavage. EpCAM (red) and DAPI (blue) staining was used to identify the intestinal epithelial layer. The experiment was repeated two times with similar results.

(B) qRT-PCR analysis of SAA1/2 expression in the epithelial fraction of terminal ileum from Villin-Cre mice sufficient or deficient for epithelial Stat3 at 4 days after SFB colonization. Data were normalized using Gapdh.

(C) Kinetics of SFB colonization post-oral gavage of SFB^e or JAX fecal contents. SFB genomic 16S in terminal ileum was quantified by qPCR analysis. Copy number of the SFB genomic 16S was normalized to host *Il-23r* genomic DNA (JAX; n = 4. SFB^e; n = 7).

(D and E) Kinetics of IL-22 (D) and SAA1/2 (E) mRNA expression post-SFB^e gavage. (D) Level of IL-22 mRNA from isolated ileal lamina propria cells was assessed (JAX; n = 7. SFB^e; n = 8 in each time points). (E) Level of SAA1/2 mRNA from epithelial fraction in ileum (JAX; n = 4. SFB^e; n = 7). Graphs (C–E) show accumulated data from two or three independent experiments as mean with SD. Day 0 samples were from animals not gavaged with additional fecal samples.

(F) SAA1/2 mRNA expression, determined by qRT-PCR, in epithelial fraction of terminal ileum from IL-22 sufficient or deficient mice at 4 days after SFB gavage (JAX gavaged: n = 10, SFB^e gavaged IL-22^{+/+} and ^{+/−}: n = 5, SFB^e gavaged IL-22^{−/−}: n = 5).

(G) SAA1/2 mRNA expression in small intestinal epithelial cell organoid cultures following rIL-22 treatment.

All qRT-PCR experiments were repeated two times with similar results, and data were normalized to GAPDH mRNA. Each dot represents a single mouse, and the orange bar indicates average relative gene expression of each group. N.S., not significant (p > 0.05); *p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001. See also Figure S4.

we observed specific SFB colonization of the ileum by 24 hr post-SFB gavage. Within 5 days, SFB antigen-presentation occurred in the MLN, where a significant

number of RORγt⁺ Th17 cells were first observed. By day 7, RORγt⁺ Th17 cells accumulated in the lamina propria of the intestine. During the early phase of this commensal-induced host response, IL-22 and SAA1/2 mRNA were acutely upregulated in the ileum within 3 days. Our studies revealed that activation of epithelial SAAs involved an SFB-initiated signaling pathway requiring IL-23 and ILC3-derived IL-22 (Figure S6). A similar IL-22-dependent pathway was recently described in SFB induction of *Fut2*, an enzyme that catalyzes fucosylation of glycans, thus providing nutritional substrates for commensals

DISCUSSION

In this report, we describe a two-step process for the functional differentiation of intestinal Th17 cells following colonization of mice with SFB. The first step is the priming and polarization of antigen-specific CD4⁺ T cells in the tissue-draining lymph nodes, resulting in a poised state marked by the expression of RORγt, and the second is the activation of a cytokine gene expression program in a tissue microenvironment in which epithelial cell-derived factors SAA1 and SAA2 act on poised cells. Temporally,

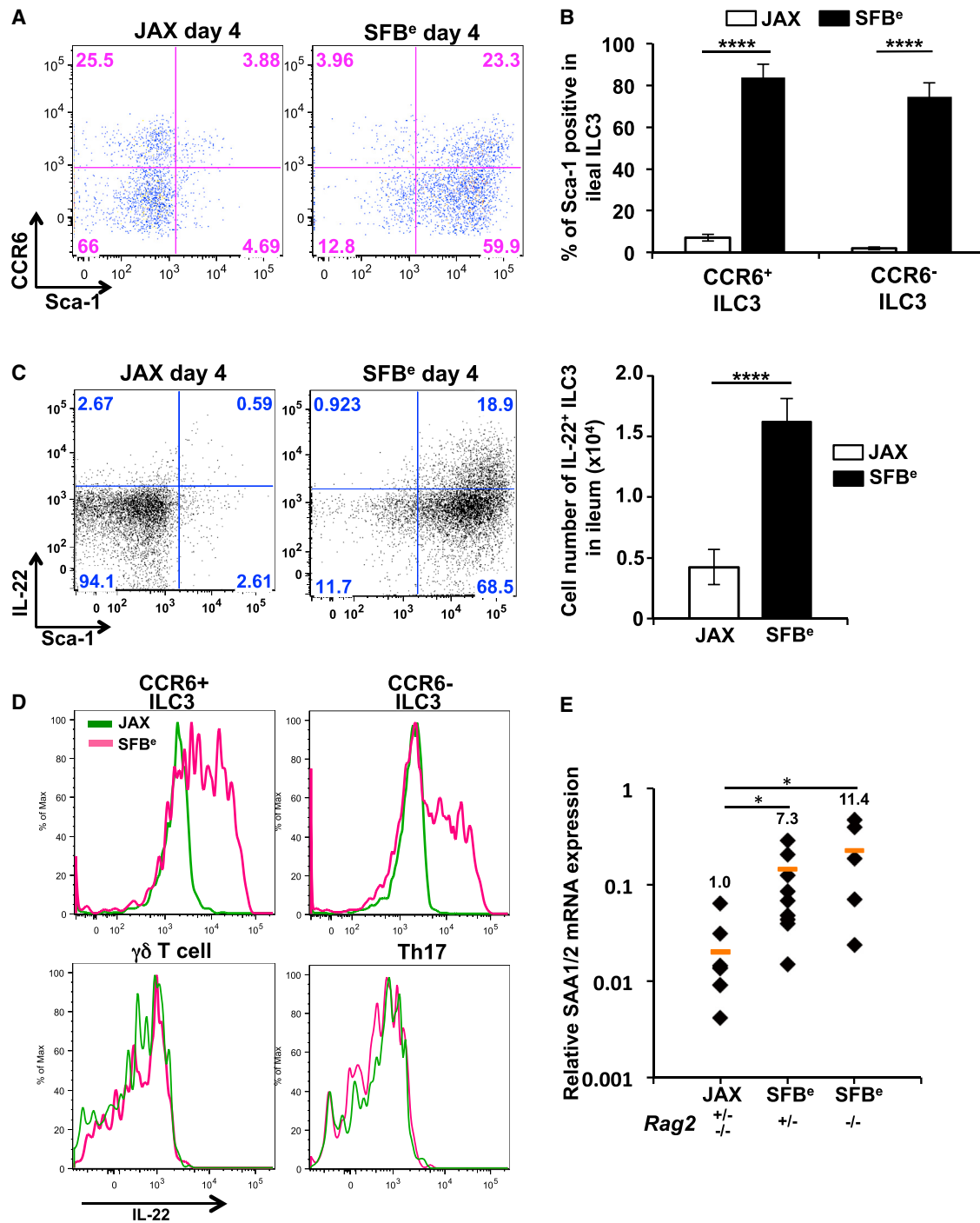


Figure 5. SFB Colonization Induces IL-22 Production by ILC3s

(A) Cell-surface expression of Sca-1 and CCR6 on ileal SILP ILC3s from mice at 4 days after oral gavage with JAX or SFB^e fecal contents. (B) Frequency of Sca-1-positive CCR6⁺ or CCR6⁻ ILC3s in the terminal ileum from JAX- or SFB^e-gavaged mice at 4 days after gavage. Data are represented as mean with SD (JAX gavaged: n = 5, SFB^e gavaged: n = 6). (C) IL-22 production in ileal SILP ILC3. Ileal LP cells from JAX- or SFB^e-gavaged mice were isolated at 4 days after gavage and cultured in vitro with GolgiStop (without PMA/Ionomycin re-stimulation) for 4 hr. IL-22-producing ILC3s were analyzed by FACS, and cell numbers were calculated (right panel). Bar graph shows accumulated data as mean with SD (JAX gavaged: n = 5, SFB^e gavaged: n = 6). (D) IL-22 production in each indicated cell type isolated from the ileal LP of SFB-free (JAX; green line) and SFB-colonized (SFB^e; magenta line) mice, analyzed at 4 days after gavage. (A–D) The experiment was repeated at least three times with similar results.

(legend continued on next page)

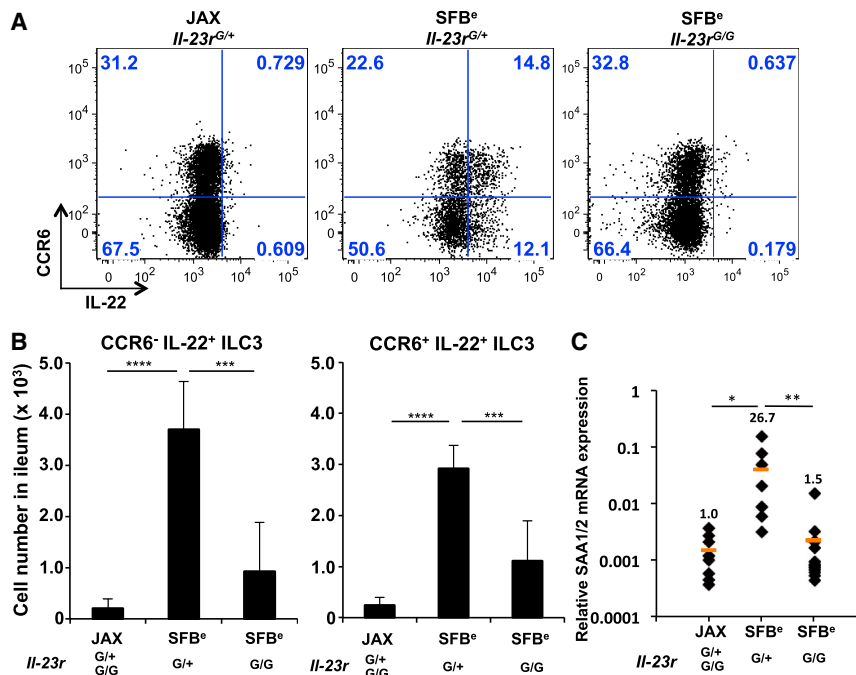


Figure 6. IL-23R Requirement for SFB-Induced Expression of IL-22 and SAA1/2 in ILC3 and Epithelial Cells, Respectively

(A) IL-22 production in CCR6⁺ and CCR6⁻ ileal LP ILC3s from *Il-23r*-sufficient or -deficient mice. (B) Cell number of IL-22-producing CCR6⁺ (left) or CCR6⁻ (right) ileal LP ILC3s in JAX- or SFB^o-gavaged mice, at 4 days after gavage. (C) Bar graphs show accumulated data from two independent experiments with a total of ten SFB-free (JAX) and five SFB-colonized (SFB^o) mice of each genotype, shown as mean with SD (C) SAA1/2 expression, analyzed by qRT-PCR, in the epithelial fraction from terminal ileum of indicated mice (JAX gavaged; n = 9, SFB^o gavaged *Il-23r*^{G/+}; n = 8, SFB^o gavaged *Il-23r*^{G/G}; n = 14). (dot) This experiment was repeated twice with similar results. Each dot represents a single mouse, and the orange bar indicates average relative gene expression of each group.

*p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001.

and contributing to barrier protection (Bry et al., 1996; Goto et al., 2014; Pickard et al., 2014).

Although we identified a requirement for the ILC3-dependent pathway during the early induction of epithelial SAA1/2, we cannot rule out that SFB attachment to the epithelial cells, which is accompanied by reorganization of the cortical actin network, contributes to or is required for SAA induction. Indeed, Atarashi et al. (2015 [this issue of Cell]) showed that rat SFB failed to attach to ileal epithelium and to induce SAA1/2 in mice, yet induced ILC3 IL-22 production. Thus, multiple signals may be required for epithelial cell induction of SAA1/2 or, alternatively, a single IL-22 signal may be sufficient, but its delivery to epithelial receptors may be compromised in mice colonized with rat SFB.

SAA in Homeostasis and Autoimmune Diseases

SFB is considered a commensal microbe, as there is no evidence that it induces pathology despite its adhesion to the epithelial surface and its induction of host adaptive and innate immune responses. By virtue of its activation of epithelial Stat3, SFB may reinforce the terminal ileal epithelial firewall, which harbors the greatest bacterial load in the small intestine (Sekirov et al., 2010). In the large intestine, Stat3 was found to mediate protection from dextran sodium sulfate (DSS)-induced colon injury in an IL-22 dependent manner. In this model, SAA induction was similarly dependent on Stat3 and required for dampening local inflammation (Eckhardt et al., 2010; Neufert et al., 2010). Recently, Hooper and colleagues showed that

SAA binds with strong affinity to vitamin A derivatives (Martineau et al., 2015), which promote mucosal homeostasis (Hall et al., 2011). Therefore, in the context of SFB colonization, the elevation of SAA1/2 and Th17 cells may be considered a homeostatic mechanism that contributes to barrier integrity in the small intestine.

SAA is significantly upregulated in the joints and serum of rheumatoid arthritis patients (O'Hara et al., 2000) and are correlated with disease progression (Chambers et al., 1983). In mouse models, Th17 cells are also known to contribute to multiple autoimmune diseases (Genovese et al., 2010; Huh et al., 2011; Lepkes et al., 2009), and whether SAA or differences in its binding of cofactors mediate these processes remain to be elucidated. As microbiota-specific helper T cells, including SFB-specific Th17 cells, circulate widely following their polarization in the gut-associated lymphoid tissues (Hand et al., 2012; Yang et al., 2014), it is possible that they contribute to autoimmune disease systemically through induction of their effector functions at sites where inflammatory mediators such as SAA are elevated.

Another possibility is that under normal circumstances SAA regulates the resolution of inflammatory responses. Th17 cells often produce IFN γ in addition to IL-17A at sites of inflammation (Ivanov et al., 2006), and this is thought to contribute to pathogenicity, but they can also give rise to IL-10-producing cells that may be important during the resolution of inflammation (Gagliani et al., 2015). Therefore, it will be important to study in detail the regulation of SAA expression in tissues during the course of autoimmune progression and whether their manipulation affects Th17-mediated pathology. Importantly, since we did not observe

(E) SAA1/2 expression, analyzed by qRT-PCR, in the epithelial fraction from terminal ileum of Rag2-sufficient or -deficient mice, at 3 days after SFB (SFB^o) or control (JAX) gavage. Data are from two independent experiments. Each symbol represents a single mouse, and the orange bar indicates average of relative gene expression in each group. Experiments were repeated two times with similar results and data were normalized by GAPDH mRNA and combined. *p < 0.05; **p < 0.01; ***p < 0.001; and ****p < 0.0001.

See also Figure S5.

a complete loss in Th17 cytokine production in SAA1/2-deficient settings during SFB induction, we speculate that other epithelial cell-derived factors likely contribute to the regulation of the Th17 effector program, which will be of great interest to study.

Innate Lymphoid Cell Regulation of Barrier Tissue Immunity

Tissue-resident ROR γ t-expressing innate lymphoid cells and “innate-like” T cells, e.g., $\gamma\delta$ 17 cells, can respond rapidly to cytokine stimulation, including IL-1 β and IL-23, by producing their own cytokines, particularly IL-22. This occurs in the absence of antigen receptor stimulation in $\gamma\delta$ 17 cells (Cua and Tato, 2010; Martin et al., 2009; Sutton et al., 2009). For ILC3s, microbiota-induced local production of IL-23, IL-1 β , and TL1a/TNFSF15 by CX₃CR1⁺ mononuclear phagocytes is required for their production of IL-22 and intestinal barrier protection functions (Longman et al., 2014). We do not know whether ROR γ t⁺ SFB-specific T cells that migrate to the intestinal lamina propria require TCR signaling for induction of effector cytokines, or whether SAA1/2, in the appropriate microenvironment, suffices for this purpose. It will be important, in future studies, to identify the receptor(s) for SAA1/2 on Th17 cells and characterize the signaling mechanism required for the cytokine response.

Our findings suggest that SAA1 and SAA2 serve as sentinels for the tissue microenvironment, enlisting T helper cells with poised Th17 functions to protect or heal the epithelial barrier. It will be important to determine whether the SAAs are unique in their ability to amplify the Th17 cell transcriptional program and whether other induced mediators similarly influence other functional T cell subsets. In this regard, IL-18 and IL-33 were recently shown to amplify Th1 effector responses during infection with *Salmonella* or *Chlamydia* (O'Donnell et al., 2014). Efforts to dissect the upstream pathways and secretory outputs that impinge on T cell effector functions at steady state and in the face of multiple challenges will create novel opportunities to modulate tissue immunity.

A critical remaining question is how SFB are sensed within the ileum. Atarashi et al. (2015) showed that adhesion of microbes to epithelium is a necessary step for Th17 differentiation. It will be important to learn whether distinct signals received by ileal epithelial cells from bound SFB are critical for Th17 cell induction and subsequent cytokine production. Characterization of such signals may provide insights into how intestinal microbes contribute to local and systemic inflammation.

EXPERIMENTAL PROCEDURES

Mouse Strains and SFB

C57BL/6 mice were obtained from Taconic Farms or the Jackson Laboratory. All transgenic animals were bred and maintained in the animal facility of the Skirball Institute (NYU School of Medicine) in specific-pathogen-free (SPF) conditions. *Il-23^{fl/fl}* mice (Awasthi et al., 2009) were provided by M. Oukka and maintained on Jackson flora. *Il-22* mutant mice were provided by Pfizer. *Saa1/2* double-knockout mice were previously described (Eckhardt et al., 2010) and were backcrossed for eight generations onto the C57BL/6 background. SFB-specific Th17-TCRTg (7B8) mice were previously described (Yang et al., 2014) and maintained on the *Il17a^{fl/fl}* reporter (JAX; C57BL/6-*Il17atm1Bcgen/J*) and Ly5.1 background (JAX; B6. SJL-Ptprca Pepcb/BoyJ). *EF1a: Lox-stop-lox-GFP-L10* mice were previously described (Stanley et al., 2013) and bred to *Cd4* Cre (Taconic), *Cd11c* Cre

(JAX), *Villin* Cre (JAX), or *Albumin* Cre (JAX) mice. *Stat3^{lox/lox}* mice (JAX, B6.129S1 *Stat3tm1Xyfu/J*) were bred to *Villin*-Cre mice. Six- to 18-week-old animals were used for experiments. Fecal pellets were provided by Y. Umesaki (Yakult) to establish a colony of SFB mono-colonized mice in our gnotobiotic animal facility. All transgenic animals were maintained in the animal facility of the Skirball Institute in accordance with protocols approved by the Institutional Animal Care and Use Committee of the NYU School of Medicine.

Acute SFB Colonization

Fecal pellets were collected from SFB-enriched *Il-23r Rag2* DKO mice (SFB⁺), SFB mono-associated mice (pure SFB), or SFB-free JAX B6 mice (JAX), respectively. Fresh fecal pellets were homogenized through a 100- μ m filter, pelleted at 3,400 rpm for 10 min, and re-suspended in PBS. Each animal was administered one-quarter pellet by oral gavage.

Statistics

Two-tailed unpaired Student's t tests were performed to compare the results using Excel and Prism. We treated less than 0.05 of p value as significant differences. *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001. All experiments were performed at least twice.

ACCESSION NUMBERS

RNA-seq data sets and Trap-seq data sets are available on GEO (GSE70599).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.061>.

AUTHOR CONTRIBUTIONS

T.S. and W.H. designed and performed most experiments and analyzed the data. J.A.H. performed experiments related to ILC3 activation, J.-Y.L. performed organoid culture studies, A.C. and S.J.G. assisted with in vivo experiments, and Y.Y. generated TCR transgenic mice and provided cells for RNA-seq analyses, W.H. and A.I.D. designed and performed Ribosome TRAP-seq studies. J.W.Z. performed studies related to IL-23R regulation of SFB-induced epithelial responses. T.S., W.H., J.A.H., and D.R.L. wrote the manuscript. D.R.L. supervised the research and participated in experimental design.

ACKNOWLEDGMENTS

We thank Sebastian Amigorena, Kenneth Cadwell, Suzan R. Schwab, and Sang V. Kim for valuable discussion, Wenjun Ouyang for providing control and anti-IL-22 neutralizing antibody, Jeffrey M. Friedman for the *EF1-lox-stop-lox-GFP-L10* mice, and Frederick De Beer for providing the *Saa* mutant mice. We thank Maria Ciofani for her contributions to the RNA-seq studies. We thank Richard M. Myers at HudsonAlpha Institute for Biotechnology for RNA-seq and TRAP-seq studies. This work was supported by fellowships from the TOYBO Bioscience foundation (T.S.), Human Frontier Science Program (T.S.), Cancer Research Institute (W.H.), and NIH T32 CA009161_Levy (W.H.), and Dale and Betty Frey Fellowship of the Damon Runyon Cancer Research Foundation 2105-12 (J.A.H.); and by the Howard Hughes Medical Institute (D.R.L.), the Helen and Martin Kimmel Center for Biology and Medicine (D.R.L.), and National Institutes of Health grant R01DK103358 (R.B. and D.R.L.).

Received: May 9, 2015

Revised: July 25, 2015

Accepted: August 17, 2015

Published: September 24, 2015

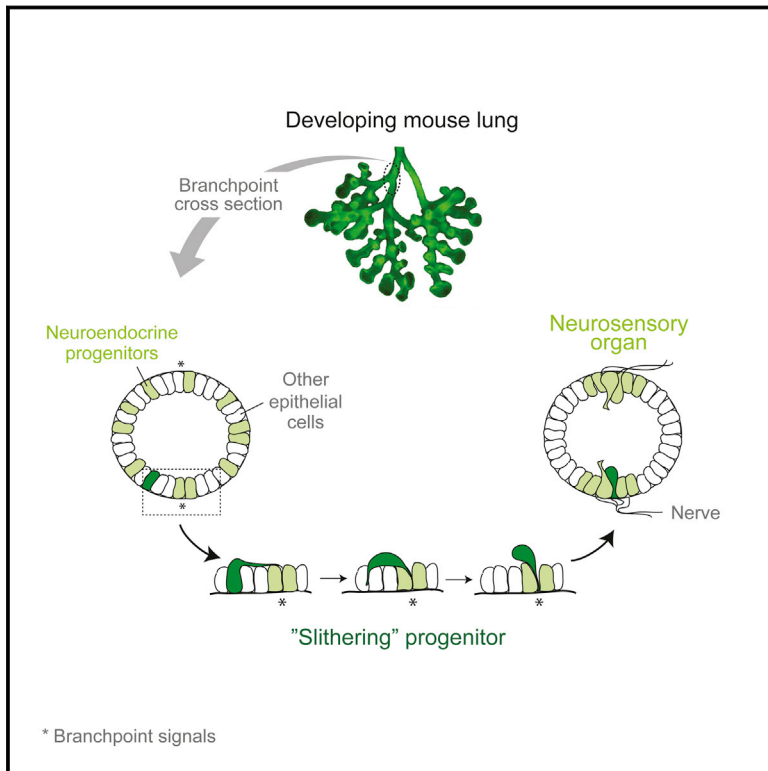
REFERENCES

- Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K., et al. (2013). Treg induction by a rationally selected mixture of *Clostridia* strains from the human microbiota. *Nature* 500, 232–236.
- Atarashi, K., Tanoue, T., Ando, M., Kamada, N., Nagano, Y., Narushima, S., Suda, W., Imaoka, A., Setoyama, H., Nagamori, T., et al. (2015). Th17 cell induction by adhesion of microbes to intestinal epithelial cells. *Cell* 163. Published online September 24, 2015. <http://dx.doi.org/10.1016/j.cell.2015.08.058>.
- Awasthi, A., Riou-Blanco, L., Jäger, A., Korn, T., Pot, C., Galileos, G., Bettelli, E., Kuchroo, V.K., and Oukka, M. (2009). Cutting edge: IL-23 receptor gfp reporter mice reveal distinct populations of IL-17-producing cells. *J. Immunol.* 182, 5904–5908.
- Bry, L., Falk, P.G., Midtvedt, T., and Gordon, J.I. (1996). A model of host-microbial interactions in an open mammalian ecosystem. *Science* 273, 1380–1383.
- Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C., and Mathis, D. (2013). A special population of regulatory T cells potentiates muscle repair. *Cell* 155, 1282–1295.
- Chambers, R.E., MacFarlane, D.G., Whicher, J.T., and Dieppe, P.A. (1983). Serum amyloid-A protein concentration in rheumatoid arthritis and its role in monitoring disease activity. *Ann. Rheum. Dis.* 42, 665–667.
- Chien, Y.H., Zeng, X., and Prinz, I. (2013). The natural and the inducible: interleukin (IL)-17-producing $\gamma\delta$ T cells. *Trends Immunol.* 34, 151–154.
- Cipolletta, D., Feuerer, M., Li, A., Kamei, N., Lee, J., Shoelson, S.E., Benoist, C., and Mathis, D. (2012). PPAR- γ is a major driver of the accumulation and phenotype of adipose tissue Treg cells. *Nature* 486, 549–553.
- Cray, C., Zaias, J., and Altman, N.H. (2009). Acute phase response in animals: a review. *Comp. Med.* 59, 517–526.
- Cua, D.J., and Tato, C.M. (2010). Innate IL-17-producing cells: the sentinels of the immune system. *Nat. Rev. Immunol.* 10, 479–489.
- Derebe, M.G., Zlatkov, C.M., Gattu, S., Ruhn, K.A., Vaishnava, S., Diehl, G.E., MacMillan, J.B., Williams, N.S., and Hooper, L.V. (2014). Serum amyloid A is a retinol binding protein that transports retinol during bacterial infection. *eLife* 3, e03206.
- Eckhardt, E.R., Witta, J., Zhong, J., Arsenescu, R., Arsenescu, V., Wang, Y., Ghoshal, S., de Beer, M.C., de Beer, F.C., and de Villiers, W.J. (2010). Intestinal epithelial serum amyloid A modulates bacterial growth in vitro and pro-inflammatory responses in mouse experimental colitis. *BMC Gastroenterol.* 10, 133.
- Eken, A., Singh, A.K., Treuting, P.M., and Oukka, M. (2014). IL-23R+ innate lymphoid cells induce colitis via interleukin-22-dependent mechanism. *Mucosal Immunol.* 7, 143–154.
- Furuzawa-Carballeda, J., Vargas-Rojas, M.I., and Cabral, A.R. (2007). Autoimmune inflammation from the Th17 perspective. *Autoimmun. Rev.* 6, 169–175.
- Gaboriau-Routhiau, V., Rakotobe, S., Lécuyer, E., Mulder, I., Lan, A., Bridonneau, C., Rochet, V., Pisi, A., De Paepe, M., Brandi, G., et al. (2009). The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* 31, 677–689.
- Gagliani, N., Vesely, M.C., Iseppon, A., Brockmann, L., Xu, H., Palm, N.W., de Zoete, M.R., Licona-Limón, P., Paiva, R.S., Ching, T., et al. (2015). Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. *Nature* 523, 221–225.
- Genovese, M.C., Van den Bosch, F., Roberson, S.A., Bojin, S., Biagini, I.M., Ryan, P., and Sloan-Lancaster, J. (2010). LY2439821, a humanized anti-interleukin-17 monoclonal antibody, in the treatment of patients with rheumatoid arthritis: A phase I randomized, double-blind, placebo-controlled, proof-of-concept study. *Arthritis Rheum.* 62, 929–939.
- Goto, Y., Obata, T., Kunisawa, J., Sato, S., Ivanov, I.I., Lamichhane, A., Takeyama, N., Kamioka, M., Sakamoto, M., Matsuki, T., et al. (2014). Innate lymphoid cells regulate intestinal epithelial cell glycosylation. *Science* 345, 1254009.
- Hall, J.A., Grainger, J.R., Spencer, S.P., and Belkaid, Y. (2011). The role of retinoic acid in tolerance and immunity. *Immunity* 35, 13–22.
- Hand, T.W., Dos Santos, L.M., Bouladoux, N., Molloy, M.J., Pagán, A.J., Pepper, M., Maynard, C.L., Elson, C.O., 3rd, and Belkaid, Y. (2012). Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. *Science* 337, 1553–1556.
- Heiman, M., Kulicke, R., Fenster, R.J., Greengard, P., and Heintz, N. (2014). Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nat. Protoc.* 9, 1282–1291.
- Honda, K., and Littman, D.R. (2012). The microbiome in infectious disease and inflammation. *Annu. Rev. Immunol.* 30, 759–795.
- Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. *Science* 336, 1268–1273.
- Huh, J.R., Leung, M.W., Huang, P., Ryan, D.A., Krout, M.R., Malapaka, R.R., Chow, J., Manel, N., Ciofani, M., Kim, S.V., et al. (2011). Digoxin and its derivatives suppress Th17 cell differentiation by antagonizing ROR γ t activity. *Nature* 472, 486–490.
- Ivanov, I.I., McKenzie, B.S., Zhou, L., Tadokoro, C.E., Lepelley, A., Lafaille, J.J., Cua, D.J., and Littman, D.R. (2006). The orphan nuclear receptor ROR γ directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell* 126, 1121–1133.
- Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., et al. (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* 139, 485–498.
- Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V.K. (2009). IL-17 and Th17 Cells. *Annu. Rev. Immunol.* 27, 485–517.
- Lathrop, S.K., Bloom, S.M., Rao, S.M., Nutsch, K., Lio, C.W., Santacruz, N., Peterson, D.A., Stappenbeck, T.S., and Hsieh, C.S. (2011). Peripheral education of the immune system by colonic commensal microbiota. *Nature* 478, 250–254.
- Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* 159, 1312–1326.
- Lee, Y.K., Menezes, J.S., Umesaki, Y., and Mazmanian, S.K. (2011). Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. *Proc. Natl. Acad. Sci. USA* 108 (Suppl 1), 4615–4622.
- Leppkes, M., Becker, C., Ivanov, I.I., Hirth, S., Wirtz, S., Neufert, C., Pouly, S., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., et al. (2009). ROR γ -expressing Th17 cells induce murine chronic intestinal inflammation via redundant effects of IL-17A and IL-17F. *Gastroenterology* 136, 257–267.
- Lewis, K.L., Caton, M.L., Bogunovic, M., Greter, M., Grajkowska, L.T., Ng, D., Klinakis, A., Charo, I.F., Jung, S., Gommerman, J.L., et al. (2011). Notch2 receptor signaling controls functional differentiation of dendritic cells in the spleen and intestine. *Immunity* 35, 780–791.
- Longman, R.S., Diehl, G.E., Victorio, D.A., Huh, J.R., Galan, C., Miraldi, E.R., Swaminath, A., Bonneau, R., Scherl, E.J., and Littman, D.R. (2014). CX₃CR1⁺ mononuclear phagocytes support colitis-associated innate lymphoid cell production of IL-22. *J. Exp. Med.* 211, 1571–1583.
- Lügering, A., Ross, M., Sieker, M., Heidemann, J., Williams, I.R., Domschke, W., and Kucharzik, T. (2010). CCR6 identifies lymphoid tissue inducer cells within cryptopatches. *Clin. Exp. Immunol.* 160, 440–449.
- Macpherson, A.J., and Harris, N.L. (2004). Interactions between commensal intestinal bacteria and the immune system. *Nat. Rev. Immunol.* 4, 478–485.
- Martin, B., Hirota, K., Cua, D.J., Stockinger, B., and Veldhoen, M. (2009). Interleukin-17-producing gammadelta T cells selectively expand in response to pathogen products and environmental signals. *Immunity* 31, 321–330.
- Martineau, A.R., James, W.Y., Hooper, R.L., Barnes, N.C., Jolliffe, D.A., Greiller, C.L., Islam, K., McLaughlin, D., Bhowmik, A., Timms, P.M., et al. (2015). Vitamin D3 supplementation in patients with chronic obstructive pulmonary disease (ViDICO): a multicentre, double-blind, randomised controlled trial. *Lancet Respir. Med.* 3, 120–130.

- Neufert, C., Pickert, G., Zheng, Y., Wittkopf, N., Warntjen, M., Nikolaev, A., Ouyang, W., Neurath, M.F., and Becker, C. (2010). Activation of epithelial STAT3 regulates intestinal homeostasis. *Cell Cycle* 9, 652–655.
- O'Donnell, H., Pham, O.H., Li, L.X., Atif, S.M., Lee, S.J., Ravessloot, M.M., Stolfi, J.L., Nuccio, S.P., Broz, P., Monack, D.M., et al. (2014). Toll-like receptor and inflammasome signals converge to amplify the innate bactericidal capacity of T helper 1 cells. *Immunity* 40, 213–224.
- O'Hara, R., Murphy, E.P., Whitehead, A.S., Fitzgerald, O., and Bresnihan, B. (2000). Acute-phase serum amyloid A production by rheumatoid arthritis synovial tissue. *Arthritis Res.* 2, 142–144.
- Pickard, J.M., Maurice, C.F., Kinnebrew, M.A., Abt, M.C., Schenten, D., Golovkina, T.V., Bogatyrev, S.R., Ismagilov, R.F., Pamer, E.G., Turnbaugh, P.J., and Chervonsky, A.V. (2014). Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature* 514, 638–641.
- Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., and Mazmanian, S.K. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science* 332, 974–977.
- Rutz, S., Wang, X., and Ouyang, W. (2014). The IL-20 subfamily of cytokines—from host defence to tissue homeostasis. *Nat. Rev. Immunol.* 14, 783–795.
- Sato, T., and Clevers, H. (2013). Primary mouse small intestinal epithelial cell cultures. *Methods Mol. Biol.* 945, 319–328.
- Sawa, S., Lochner, M., Satoh-Takayama, N., Dulauroy, S., Bérard, M., Kleinschek, M., Cua, D., Di Santo, J.P., and Eberl, G. (2011). ROR γ t⁺ innate lymphoid cells regulate intestinal homeostasis by integrating negative signals from the symbiotic microbiota. *Nat. Immunol.* 12, 320–326.
- Sekirov, I., Russell, S.L., Antunes, L.C., and Finlay, B.B. (2010). Gut microbiota in health and disease. *Physiol. Rev.* 90, 859–904.
- Spits, H., and Di Santo, J.P. (2011). The expanding family of innate lymphoid cells: regulators and effectors of immunity and tissue remodeling. *Nat. Immunol.* 12, 21–27.
- Stanley, S., Domingos, A.I., Kelly, L., Garfield, A., Damanpour, S., Heisler, L., and Friedman, J. (2013). Profiling of Glucose-Sensing Neurons Reveals that GHRH Neurons Are Activated by Hypoglycemia. *Cell Metab.* 18, 596–607.
- Stockinger, B., and Veldhoen, M. (2007). Differentiation and function of Th17 T cells. *Curr. Opin. Immunol.* 19, 281–286.
- Sutton, C.E., Lalor, S.J., Sweeney, C.M., Brereton, C.F., Lavelle, E.C., and Mills, K.H. (2009). Interleukin-1 and IL-23 induce innate IL-17 production from gammadelta T cells, amplifying Th17 responses and autoimmunity. *Immunity* 31, 331–341.
- Uhlir, C.M., and Whitehead, A.S. (1999). Serum amyloid A, the major vertebrate acute-phase reactant. *Eur. J. Biochem.* 265, 501–523.
- Virgin, H.W. (2014). The virome in mammalian physiology and disease. *Cell* 157, 142–150.
- Weaver, C.T., Hatton, R.D., Mangan, P.R., and Harrington, L.E. (2007). IL-17 family cytokines and the expanding diversity of effector T cell lineages. *Annu. Rev. Immunol.* 25, 821–852.
- Wu, H.J., Ivanov, I.I., Darce, J., Hattori, K., Shima, T., Umesaki, Y., Littman, D.R., Benoist, C., and Mathis, D. (2010). Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* 32, 815–827.
- Yang, Y., Torchinsky, M.B., Gobert, M., Xiong, H., Xu, M., Linehan, J.L., Alonzo, F., Ng, C., Chen, A., Lin, X., et al. (2014). Focused specificity of intestinal TH17 cells towards commensal bacterial antigens. *Nature* 510, 152–156.

Formation of a Neurosensory Organ by Epithelial Cell Slithering

Graphical Abstract



Authors

Christin S. Kuo, Mark A. Krasnow

Correspondence

krasnow@stanford.edu

In Brief

A form of migration named "slithering," where cells undergo transient EMT, allows rapid assembly of pulmonary neurosensory organs and may explain why lung cancers arising from neuroendocrine cells are highly metastatic.

Highlights

- Neuroendocrine cells form clusters at stereotyped positions at branch junctions
- Progenitors are selected randomly in epithelium but rapidly sort from other cells
- Sorting occurs by directed migration of progenitors to branch junctions
- Migrating cells undergo transient EMT, crawling over neighbor cells to the target

Formation of a Neurosensory Organ by Epithelial Cell Slithering

Christin S. Kuo^{1,2,3} and Mark A. Krasnow^{1,3,*}

¹Department of Biochemistry

²Department of Pediatrics

³Howard Hughes Medical Institute

Stanford University School of Medicine, Stanford, CA 94305-5307, USA

*Correspondence: krasnow@stanford.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.021>

SUMMARY

Epithelial cells are normally stably anchored, maintaining their relative positions and association with the basement membrane. Developmental rearrangements occur through cell intercalation, and cells can delaminate during epithelial-mesenchymal transitions and metastasis. We mapped the formation of lung neuroepithelial bodies (NEBs), innervated clusters of neuroendocrine/neurosensory cells within the bronchial epithelium, revealing a targeted mode of cell migration that we named “slithering,” in which cells transiently lose epithelial character but remain associated with the membrane while traversing neighboring epithelial cells to reach cluster sites. Immunostaining, lineage tracing, clonal analysis, and live imaging showed that NEB progenitors, initially distributed randomly, downregulate adhesion and polarity proteins, crawling over and between neighboring cells to converge at diametrically opposed positions at bronchial branchpoints, where they reestablish epithelial structure and express neuroendocrine genes. There is little accompanying progenitor proliferation or apoptosis. Activation of the slithering program may explain why lung cancers arising from neuroendocrine cells are highly metastatic.

INTRODUCTION

Epithelia are sheets of cells that line and protect the body and internal organs, and the polarized cells that comprise them play important roles in absorption, secretion, and sensation. Epithelial cells are normally tightly attached to one another through specialized junctions and adhesion proteins along their lateral surface and anchored to the basement membrane at their basal surface. Although epithelial sheets can grow and change shape, the constituent cells typically maintain their relative positions. When cells in an epithelial monolayer have been found to rearrange, as in *Drosophila* germband elongation (Irvine and Wieschaus, 1994) and wing morphogenesis (Aigouy et al., 2010),

they do so conservatively by cell intercalation, in which cells shrink lateral junctions with some neighboring cells while expanding lateral junctions with others, exchanging positions while maintaining their polarized structure and the integrity of the epithelium (Bertet et al., 2004; Blankenship et al., 2006; Guillot and Lecuit, 2013). Here, we describe a very different mode of epithelial cell rearrangement that results in homotypic sorting (Krens and Heisenberg, 2011) of a specialized cell type, discovered in our dissection of pulmonary neuroendocrine (NE) cell development in mice.

Pulmonary NE cells are one of the most interesting but least understood cell types in the lung. They are distributed throughout the bronchial epithelium, interspersed among secretory club (Clara) cells and ciliated cells, the two major airway epithelial cell types (Rock and Hogan, 2011). Like other neuroendocrine cells in the body, they were originally identified by their secretory dense-core vesicles (Feyrter, 1954) that contain signaling molecules and bioactive peptides, including serotonin and calcitonin gene-related peptide (CGRP). Although some pulmonary NE cells are distributed randomly in the airway epithelium, others are organized into clusters called neuroendocrine or neuroepithelial bodies (NEBs) that are highly innervated (Brouns et al., 2009; Lauweryns and Peuskens, 1972), forming synaptic contacts with afferent and efferent nerve fibers (Lauweryns and Van Lommel, 1987). NE cells can be activated by a variety of stimuli and are thought to monitor diverse aspects of lung physiology including oxygen, chemical, and mechanical changes (Cutz et al., 2013). In addition to these sensory and neurosecretory functions, NE cells have a stem cell function that helps replenish the bronchial epithelium following severe injury (Guha et al., 2012; Reynolds et al., 2000; Song et al., 2012). They are also the initiating cells of small cell lung cancer (Park et al., 2011; Song et al., 2012; Sutherland et al., 2011), a highly metastatic and the most deadly form of lung cancer (van Meerbeeck et al., 2011). Excess or altered distribution of NE cells are also found in a variety of serious, but poorly understood, lung diseases including sudden infant death syndrome (SIDS) (Cutz et al., 2007), bronchopulmonary dysplasia (BPD) (Gillan and Cutz, 1993), and neuroendocrine hyperplasia (Aguayo et al., 1992; Deterding et al., 2005).

To provide a foundation for a genetic dissection of the development, function, and diseases of pulmonary NE cells, we first mapped their locations in mice and found that NEBs are located at stereotyped positions. We then probed NEB development by

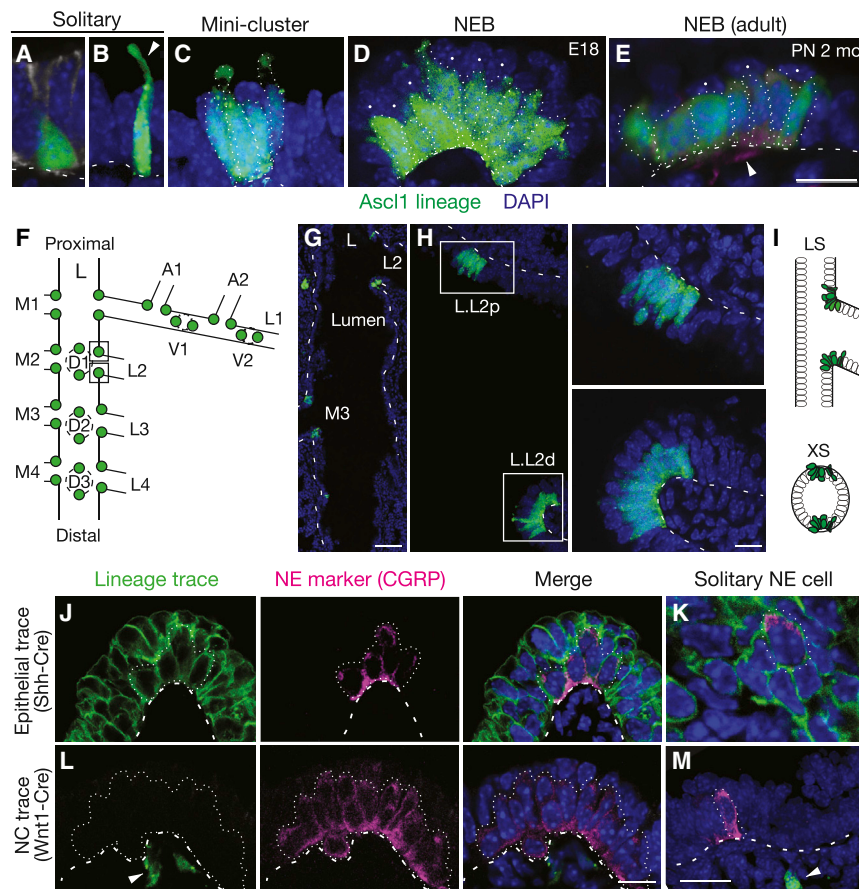


Figure 1. Mapping and Lineage Tracing of Pulmonary Neuroendocrine Cells

(A–E) Confocal images of NE cells in bronchial epithelium of embryonic day (E) 18 (A–D) or adult (postnatal 2 month; E) *Ascl1^{CreER/+}; Rosa26^{ZsGreen/+}* mouse induced with tamoxifen at E14 to label NE cells with ZsGreen (green). See also Figure S1. (A and B) Solitary NE cells with pyramidal (A) and slender (B) morphologies. Arrowhead, thin projection to lumen. (C) Mini-cluster of four NE cells. (D) Neuroepithelial body (NEB), a large cluster of ~25 NE cells (12 visible in optical plane). (E) Adult NEB. Cells are more uniformly columnar. Note non-NE cells (nuclei marked by dots) superficial to NE cells in NEBs (D and E). Sample was also stained with PGP9.5 (red) to confirm NE identity and show innervating fibers (arrowhead). Dotted lines, individual NE cells in clusters; dashed lines, basement membrane location determined by co-staining with E-cadherin or laminin $\gamma 1$. Scale bar, 10 μ m.

(F–I) Stereotyped locations of NEBs at airway branchpoints. (F) Left primary bronchus (L) and some daughter (L.L1, L.L2, ...) and granddaughter (L.L1.A1, L.L1.A2, ...) branches showing positions of NEBs at E18. (G) Low magnification confocal image of left primary bronchus of E18 *Ascl1^{CreER/+}; Rosa26^{ZsGreen/+}* mouse. Note pairs of NEBs (green) at indicated branch points (L.L2, L.M3), with one NEB located proximal (p) and the other distal (d) along bronchial tree. Scale bar, 100 μ m. (H) Longitudinal optical section through base of L.L2 showing L.L2p and L.L2d NEBs. Close-ups of NEBs (boxed) are shown at right. Scale bar, 5 μ m. (I) Schematic longitudinal section (LS) and cross section (XS) showing diametrically opposite positions of L.L2 NEBs at branch base. See also Figure S2.

(J–M) Lineage tracing of NE cell origins. Confocal images of NEBs (J and L) and solitary NE cells (K and M) at E18 from *Shh^{Cre/+}; Rosa26^{mTmG/+}* mouse to label bronchial epithelial lineage (J and K) or from *Wnt1-Cre; Rosa26^{ZsGreen/+}* mouse to label neural crest lineage (L and M). Lungs were co-stained for lineage tag (green) and NE marker CGRP (magenta). All NEB and solitary NE cells express bronchial trace and none express neural crest trace. Other epithelial cells also express bronchial trace, and sub-epithelial neurons (arrowheads) express Wnt1 trace. Dots, outline of NEBs or individual NE cells; dashes, basement membrane. Scale bar, 10 μ m.

immunostaining, lineage tracing, and imaging of developmental intermediates and found that although progenitors are initially distributed randomly throughout the epithelium, they rapidly resolve into clusters. We show that clusters do not form by progenitor proliferation, but by a targeted mechanism of epithelial cell rearrangement in which progenitors transiently lose epithelial character as they “slither” over and around neighboring cells and converge at cluster sites.

RESULTS

Mapping the Origin and Distribution of NEBs

Pulmonary NE cells are distributed sparsely throughout the bronchial epithelium both as solitary cells and clusters. Most mature NE cells are typical columnar epithelial cells, however, some have distinct morphologies such as short “pyramidal” cells that do not reach the surface or slender cells with a thin luminal projection (Figures 1A and 1B). Clusters are either small, typically with two to five NE cells, which we call “mini-clusters” (Figure 1C), or are larger clusters typically containing 20–30 NE cells

(Figures 1D and 1E). The terms “neuroepithelial body” (NEB) and “neuroendocrine body” often refer to all NE cell clusters, but here, we reserve the term NEB for large clusters. NE clusters in other species are innervated (Lauweryns and Peuskens, 1972; Scheuermann, 1987), and we found by immunostaining for neurites with PGP9.5 (Uchl1) that mouse NEBs are extensively innervated and mini-clusters show some innervation, whereas solitary NE cells do not appear to be innervated.

We mapped the locations of NE cells in the mouse bronchial tree in late stage embryonic lungs (embryonic days E16–E18) using the airway branch lineage (Metzger et al., 2008). NE cells were detected by immunostaining for the proneural bHLH transcription factor *Ascl1* (Mash1) or cytoplasmic NE marker PGP9.5, or using an *Ascl1^{CreER}* lineage trace, which within the bronchial epithelium labels exclusively the NE lineage (Figure S1). We analyzed serial sections of left lung lobes co-stained for E-cadherin to visualize airway epithelium. Solitary NE cells and mini-clusters were scattered throughout the bronchial epithelium in no obvious pattern. The pattern of NEBs, however, was highly stereotyped. NEBs localized exclusively to airway branchpoints

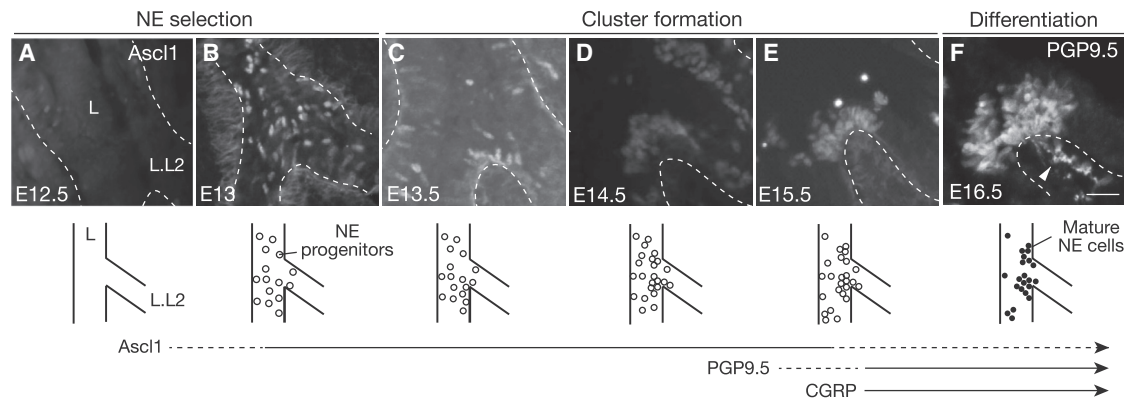


Figure 2. Cellular Events in NEB Formation

Longitudinal optical sections and schematics of base of L.L2 bronchus at stages indicated in wild-type lungs stained for NE progenitor marker Ascl1 (A–E) or mature marker PGP9.5 (F). Dashed lines, basement membrane. Three phases of NEB development are indicated: NE cell selection in bronchial epithelium, as indicated by expression of Ascl1 (open circles); NE cell cluster formation at NEB sites; and differentiation marked by expression of PGP9.5 (filled circles) and CGRP (see Figure S3F). Timing of Ascl1, PGP9.5, and CGRP expression is indicated below. PGP9.5-positive nerve fibers (arrowhead, F) extend toward and begin to ramify on L.L2 NEB at E16.5. Scale bar, 25 μ m. See also Figure S3.

(Figures 1F and 1G). Although none were found at the origin of the trachea or right and left primary bronchi, at every branchpoint examined beyond this a pair of NEBs was present (Figures 1F, 1G, and S2). NEBs were located at diametrically opposite positions at branch junctions, with one at the most proximal position and the other the most distal (Figures 1H and 1I). Hence, we designate each NEB by the name of the branch it originates, plus “p” or “d” to distinguish proximal and distal partners. The proximal NEB marks the most obtuse angle of the branch junction, whereas the distal NEB lies at the most acute angle. The same pattern was observed at every branchpoint examined in other lobes (Figure S2) ($n = 5$ lungs, 70 branchpoints). Thus, NEBs form at stereotyped, diametrically opposite positions at secondary airway branchpoints and beyond, implying there is a localized signal directing NEB formation at each of these sites.

NE cells in other organs arise either from invading neural crest cells that colonize the organ or by differentiation of resident epithelial cells (Anderson and Axel, 1986; Gu et al., 2002). Initial lineage trace studies indicated that at least some pulmonary NE cells arise from the epithelium (Song et al., 2012). To determine if NEBs arise from the epithelium, we used a *Sonic hedgehog*^{Cre} (*Shh*^{Cre}) knock-in allele (Harris et al., 2006) and a Cre recombinase reporter to label and lineage trace developing airway epithelial progenitors several days before NE progenitors are detected. At E18.5, all NEBs as well as mini-clusters and solitary pulmonary NE cells expressed the *Shh*^{Cre} lineage trace ($n = 500$ NE cells scored) (Figures 1J and 1K). In a similar experiment using a *Wnt1*-Cre transgene to lineage trace developing neural crest cells, no NE cells at E18.5 expressed the lineage trace ($n = 500$) (Figure 1L). We conclude that all pulmonary NE cells, including NEBs, arise from the *Shh* epithelial lineage, not from neural crest. Thus, both solitary and clustered pulmonary NE cells share the same origin as other airway epithelial cells.

Three Phases of NEB Development

To elucidate the cellular events and timing of NEB formation, we mapped the development of specific NEBs beginning at E12,

before NE progenitors are detected, through E16 when NEBs achieve their characteristic structure. We used Ascl1, the earliest known marker of NE cells (Ito et al., 2000), to visualize progenitors and early events and PGP9.5 and CGRP, canonical markers of mature NE cells (Polak et al., 1993), to visualize later steps. We analyzed five serially sectioned wild-type CD-1 lungs at each of six stages between E12–E16, initially focusing on the base of branch L.L2 where NEBs L.L2p and L.L2d form (Figure 2). The earliest NE progenitors at these sites were detected at late E12.5–E13, when a few scattered cells in the epithelium began to express Ascl1 (Figure 2A). At E13, more densely distributed but still solitary NE progenitors were present in a “salt and pepper” pattern (Figure 2B). By late E13.5–E14, small groups of NE progenitors were detected at L.L2p and L.L2d and also scattered around the region (Figure 2C). By E14.5, within 24–48 hr of detection of the first NE progenitors, numerous small clusters of NE cells began to appear throughout the region, interspersed with solitary cells (Figures 2D and S3B). At E15.5, large clusters of NE cells were detected, but exclusively at the positions of NEBs L.L2p and L.L2d (Figure 2E). By E15.5–E16, NE cells in NEBs L.L2p and L.L2d expressed PGP9.5 and CGRP, as did solitary NE cells and mini-clusters in the region (Figures 2F, S3E, and S3F). This analysis subdivided NEB formation into three developmental phases: NE cell selection (E12.5–E13.5), cluster formation (E13.5–E15.5), and differentiation (E15.5 onward). During differentiation, PGP9.5-positive nerve fibers extend toward and begin to ramify on the NEB (Figures 2F and S3G–S3I).

The same series of cellular events and developmental phases was observed at other positions where NEB formation was mapped, including L.L3p/L.L3d, L.L4p/L.L4d, and L.L5p/L.L5d, all located at the base of secondary branches along the left primary bronchus. However, the process initiates slightly later at the more distal positions, delayed by ~0.5 day for L.L3, 1 day for L.L4, and 1.5 days for L.L5 NEBs. This is best appreciated in images capturing all four positions in the same lung, when the most distal branch (L.L5) has not yet initiated NE development, L.L4 is

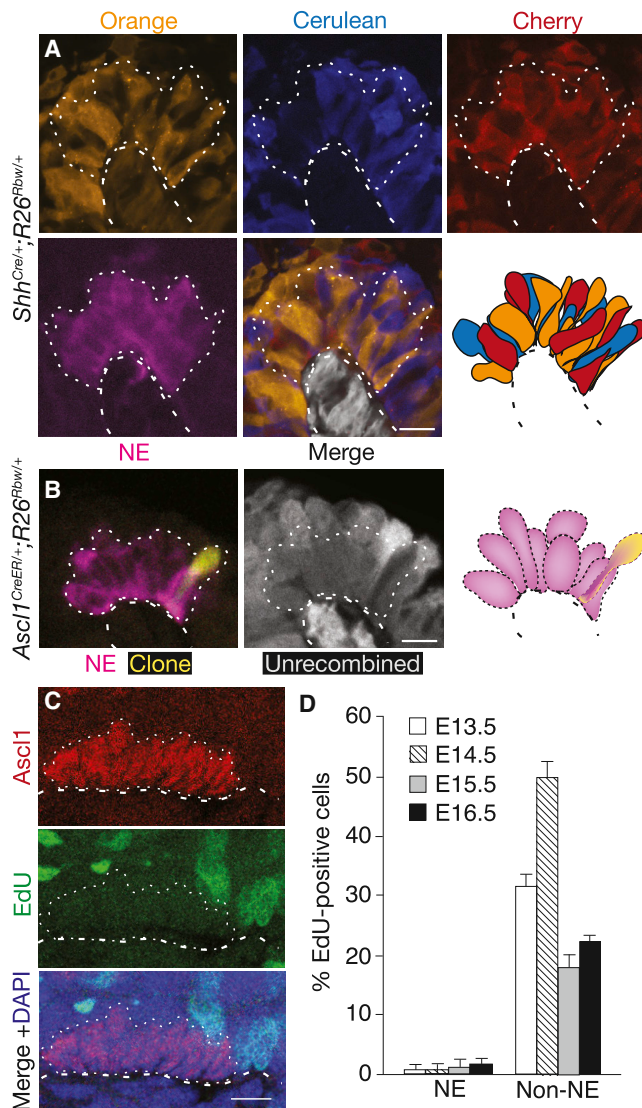


Figure 3. Analysis of Proliferation during NEB Formation

(A) Close-up of NEB in *Shh*^{Cre/+};*Rosa26*^{Rbw/+} mouse lung in which entire bronchial epithelium was permanently labeled with one of the Rainbow reporters (Orange, Cerulean, or Cherry) before NE selection then co-stained for CGRP (magenta) at E18 following NEB formation. NEB (dotted outline) is composed of cells labeled with each of the reporters, thus it did not arise by proliferation of a single progenitor. Epithelial cells neighboring the NEB are not shown in schematic; cells below basement membrane (dashed line) express the constitutive (expressed in absence of Cre) GFP reporter (pseudo-colored gray). Scale bar, 10 μ m. See also Table S1.

(B) Close-up of a NEB in E18 *Ascl1*^{CreER/+};*Rosa26*^{Rbw/+} mouse lung in which NE progenitors were sparsely labeled by injection of tamoxifen (4 mg) early in NE development (E11.5) then immunostained for CGRP (magenta) 6 days later following NEB formation. Note one labeled cell (pseudocolored yellow) in NEB, indicating that labeled progenitor did not proliferate. Cells without Cre activity ("Unrecombined") express the GFP reporter (pseudo-colored gray in middle panel); note absence of GFP in clone (yellow cell). Only cells in NEB (dotted lines) are shown in schematic. Scale bar, 10 μ m. See also Tables S2A–S2C.

(C) Confocal sagittal section through cluster of NE cells during NEB formation in E14 wild-type embryo in which EdU was injected 2 hr before harvesting to label dividing cells, then analyzed after co-staining for EdU (green) and

undergoing NE cell selection, L.L3 has begun NE clustering, and L.L2 is in the NE differentiation phase (Figures S3A and S3B). Thus, NEBs form in three phases—NE selection, cluster formation, and differentiation—that initiate in a proximal-to-distal sequence separated by approximately a half day between branches along the left primary bronchus. Despite differences in the timing of NEB program initiation along the proximal-distal axis, the differentiation phase begins in near synchrony for all NEBs and other NE cells (Figures S3G–S3I), suggesting that this final phase is triggered by a more global signal independent of earlier steps in the program.

Cluster Formation Does Not Occur by Proliferation of NE Precursors

NE and neurosensory cells in many tissues are arranged in clusters, but how clusters form is only understood for *Drosophila* and other insect sensory bristles. Each *Drosophila* bristle arises by proliferation of a single progenitor expressing Achaete-Scute proteins (founders of the *Ascl1* gene family), and the clonal cell cluster differentiates into bristle sensory and support cells (Cubas et al., 1991; Lawrence, 1966). To determine if NEBs form by clonal proliferation of an epithelial progenitor, we labeled all airway epithelial progenitors prior to NEB formation using *Shh*^{Cre} and the "Rainbow" (*Rosa26*^{Rbw}) Cre-reporter (Rinkevich et al., 2011), which stochastically and permanently labels each progenitor and any of its daughter cells with one of three different fluorescent proteins. If NEBs form like *Drosophila* sensory bristles, each mature NEB should be monoclonal and hence composed of cells of a single Rainbow color. We found instead that all NEBs examined at E17.5–E18 ($n > 100$), including those examined closely by confocal microscopy to distinguish individual NE cells, were multi-colored (Figure 3A; Table S1). A similar result obtained for NE miniclusters. Thus, NEBs and miniclusters arise from multiple founder cells, not clonal proliferation of a single NE progenitor.

Although the above results rule out NEB formation by proliferation of a single progenitor, they do not exclude the possibility that NEBs are polyclonal, arising by proliferation of multiple NE progenitors. We therefore carried out two types of experiments to investigate the role of progenitor proliferation in NEB formation. We first used an inducible Cre recombinase (Cre-ER) expressed from the endogenous *Ascl1* promoter (*Ascl1*^{CreER}) (Kim et al., 2011), and an early (~E11.5–E12) dose of the inducer tamoxifen, to sparsely label individual NE progenitors with one of the Rainbow colors. When mature NEBs and miniclusters were examined 6 days later, few clusters were labeled, and of those that were, most contained just a single labeled cell (30 of 39 labeled clusters) (Figure 3B). In those with more than one labeled

Ascl1 (red). No NE cells, but many surrounding epithelial cells, are labeled with EdU.

(D) Quantification of EdU-positive NE and non-NE airway epithelial cells at indicated ages. Values shown are percent \pm SE for proportions ($n = 134, 119, 78, 130$ [NE cells], $n = 520, 322, 326, 930$ [non-NE cells] at E13.5, 14.5, 15.5, 16.5, respectively) scored in two embryos for each cell type scored at each age. Note few EdU-positive NE cells, indicating progenitors exit cell cycle at or just after selection. Similar results were obtained for phospho-histone H3 analysis. Scale bar, 10 μ m.

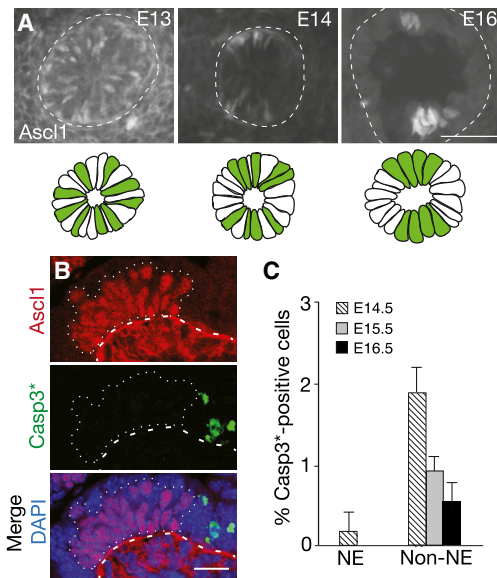


Figure 4. Local Clearing of NE Cell Precursors during Cluster Formation

(A) Cross-sections through base of bronchial branch L.V2 at indicated ages of wild-type embryos immunostained for *Ascl1* (white; green in schematics below) to show NE progenitors forming NEBs L.V2p and L.V2d. Progenitors are initially scattered in “salt and pepper” pattern but, as NEBs form, surrounding regions clear of NE cells. Scale bar, 10 μ m. (B) Sagittal section through developing NEB (dotted outline) in E14 lung co-stained for *Ascl1* (red) and activated (cleaved) Caspase 3 (Casp3*, green) to detect apoptotic cells. Note Casp3*-positive cells near, but not within, NEB (dotted outline). *Ascl1* signal below basement membrane (dashed line) at E13 and E14 is non-specific background. Scale bar, 10 μ m. (C) Quantification of Casp3*-positive NE and non-NE cells in bronchial epithelium at indicated stages. Data shown are percent \pm SE for proportions; $n = 589, 406, 322$ (NE cells), 1,437, 1,534, 1,122 (non-NE cells) scored at E14.5, 15.5, 16.5, respectively, from two embryos at each stage.

cell, the cells were usually different colors, in the ratio expected for independent labeling events (Table S2). These clonal labeling results imply that there is little or no proliferation of NE progenitors once they express the progenitor marker *Ascl1*. A similar conclusion obtained by pulse-labeling with deoxyribonucleotide analog EdU from E12–E15 to detect dividing NE cells. Although EdU-labeled cells were readily detected in surrounding (non-neuroendocrine) epithelial cells and mesenchyme, few labeled NE cells were detected (Figures 3C and 3D). The only labeled NE cells (<1% of NE cells) were solitary NE progenitors in the most distal regions of the bronchial tree just initiating *Ascl1* expression. We conclude that NE progenitors cease dividing at or shortly after initiation of *Ascl1* expression, and proliferation of progenitors does not contribute significantly to cluster formation.

Local Clearing of NE Cell Precursors without Apoptosis

Inspection of the airway epithelium immediately surrounding positions where NEBs form showed local “clearing” of NE progenitors: progenitors were initially evenly distributed in a salt and pepper pattern like other regions, but labeled cells were lost as NEBs formed nearby (Figure 4A). To determine if clearing

was due to programmed cell death, we examined the regions for expression of the apoptosis marker, activated Caspase-3. Although activated Caspase-3 was detected in other developing cell types, little or none was detected in *Ascl1*-expressing NE progenitors (Figures 4B and 4C). Cell counts also showed no decline in NE progenitors during cluster formation (Figures S3C and S3D). Thus, apoptosis does not contribute to clearing of NE progenitors from regions surrounding NEB formation.

Sparse Labeling Reveals a Series of Migratory NE Intermediates

Local clearing of progenitors and declining number of solitary NE cells as the number of clustered cells and size of the clusters increased (Figures S3C and S3D) suggested that NEB formation might occur by local sorting of NE progenitors. To identify and visualize developmental intermediates, we sparsely labeled NE progenitors at E11.5 or E12.5 using *Ascl1*^{CreER} and limiting tamoxifen in conjunction with the *Rosa26*^{ZsGreen} reporter that robustly labels NE progenitors, allowing visualization of the full structure of individual intermediates (Figures S4 and S5). Examination of labeled progenitors 2–4 days later (Figure 5A) revealed a collection of developmental intermediates with striking migratory morphologies, including long cytoplasmic extensions, fibroblast-like shapes oriented orthogonal to neighboring epithelial cells—some fully detached from the basement membrane—and cells spiraling around each other and converging at the basement membrane (Figures 5B–5F). NE cells with migratory structures were abundant in regions surrounding sites of NEB formation, and they were transient intermediates detected only during NEB formation (labeling between E12 and E13 but not at E14 or later) that “chased” into mature NEBs in the 36 hr after labeling. The migratory morphologies were specific to the NE lineage as other epithelial cells labeled during this period using *FoxA2*^{CreER} showed the typical structure of monolayer or pseudostratified epithelial cells (Figure S6).

Analysis of over 180 labeled NE intermediates (Table S3) by confocal microscopy distinguished five morphologic classes. Class 1 cells were the earliest and simplest intermediates: solitary NE cell progenitors lacking extensions or unusual morphologic features (Figure 5B). They are early progenitors, presumably before migration has occurred, as they were the only labeled cells detected during the earliest pulse-chase interval (E11.5–E13.5). Class 5 intermediates were similar in structure to Class 1 cells, each showing normal epithelial structure and orientation, except class 5 cells were found in clusters (Figure 5F). Class 5 cells are the most advanced intermediates because they were the only type detected at E16.5 and later (Figure 5G). The other three classes had migratory morphologies and were the most striking and unusual structures. Class 2 intermediates are solitary and resemble class 1 intermediates with a largely normal epithelial morphology, except they have a long (up to 30 μ m) cellular extension that reaches over or around neighboring epithelial cells, typically contacting another NE cell several cell diameters away (Figure 5C; Table S3). Class 3 intermediates have the most surprising structure, betraying their epithelial origin. They resembled fibroblasts and had protrusions and often twisted morphologies that curved over or around, and occasionally under, neighboring epithelial cells. Some had no

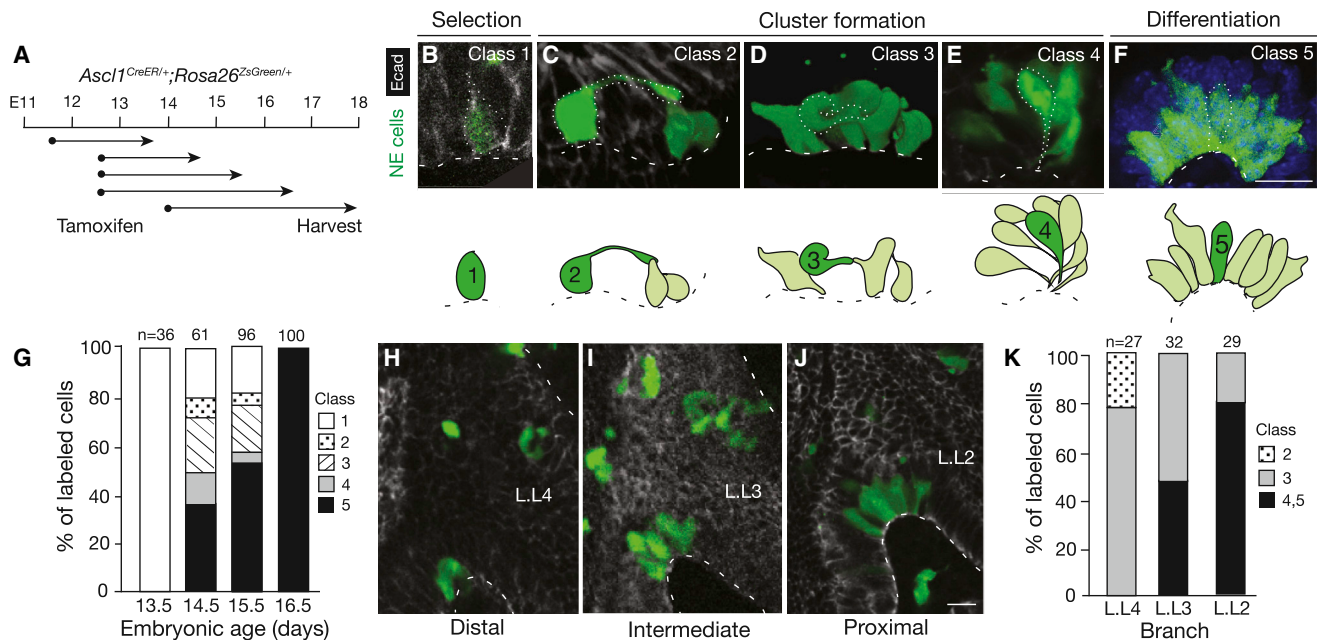


Figure 5. Sparse Labeling of Progenitors Reveals a Series of NEB Intermediates

(A) Labeling scheme. NE progenitors were permanently labeled with ZsGreen by tamoxifen injection of *Ascl1^{CreER/+};Rosa26^{ZsGreen/+}* mice at indicated ages (dots), then allowed to develop for times indicated (arrows).

(B–F) Micrographs and schematics of five classes of NE intermediates (green) labeled as above and co-stained for E-cadherin to outline epithelial cells. A representative cell of each class is outlined (dots) in micrograph and highlighted in dark green in schematic; other labeled NE cells are shown in light green in schematic. Note long (>10 μm) apical cell extension (class 2), fibroblast-like morphology and no basement membrane contact (class 3), and thin cellular extensions that converge at basement membrane (class 4). In late embryo (~E17) through adult, all NE cells contact basement membrane (class 5). Scale bar, 10 μm. See also Figures S4, S5, and Table S3.

(G) Quantification of different classes of NEB intermediates labeled as above and harvested 48 hr after tamoxifen injection at ages indicated. The changing abundance of intermediates supports their ordering in a developmental series. n, number of intermediates scored in two lungs at each stage.

(H–K) Confocal images (sagittal sections; H–J) and quantification (K) of NEB intermediates labeled as above at base of bronchi L.L4, L.L3, and L.L2 in same lung. L.L4 is most distal (immature) and L.L2 is most proximal (mature) position. The different abundance of intermediates along proximal-distal axis supports their ordering in a developmental series. n, number of intermediates scored in two embryos. Scale bar, 10 μm.

See also Figure S6.

detectable contact with the basement membrane, or contacted it only via a thin cellular extension (Figure 5D). Class 4 intermediates were associated with other NE cells, organized into nascent clusters with cell bodies localized apically and thin extensions swirling as they converged at a site on the basement membrane (Figure 5E).

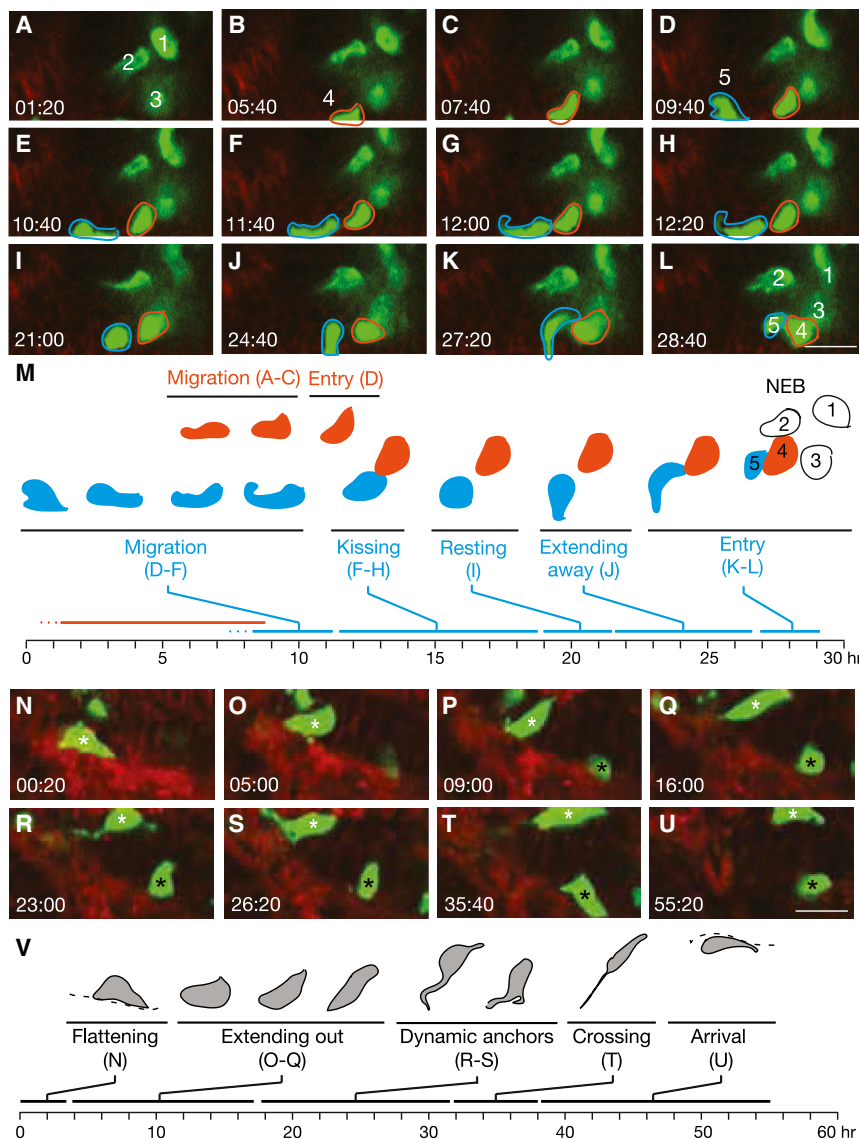
We inferred from the structures of the intermediates that they comprise a developmental series, beginning with class 1 and proceeding sequentially to class 5 (Figure 7A): NE progenitors first extend long processes and move toward their target at a branch junction, losing their epithelial structure and adopting a mesenchymal morphology as they leave their birth site. They then migrate to the target, converging there with other progenitors and re-establishing their epithelial structure and organizing into clusters. Similar intermediates were observed at sites of mini-cluster formation (Table S3), suggesting that miniclusters form by a related sorting process.

The timing and location of intermediates at sites of NEB formation support the proposed sequence. Quantification of the abundance of each type of intermediate at the same site (along the primary left lobe bronchus) but different times in NEB develop-

ment (E13.5–E16.5; Figure 5G), or at different stages in the developmental process (represented by different positions along the proximal-distal axis (L.L4, L.L3, L.L2) at the same time in development (E14) (Figures 5H–5K), supported the proposed order. Class 5 intermediates were also the only ones innervated, confirming they are the most mature. Marker expression patterns provide further support for the proposed sequence, with expression of progenitor marker *Ascl1* present in class 1 and 2 but progressively declining in class 3, 4, and 5 cells, and expression of late NE markers initiating in class 4 (PGP9.5) and 5 (CGRP) cells (Figures 2, 7A, and 7B).

Live Imaging of NE Migration during Cluster Formation

To investigate dynamics of progenitor migration and sorting, we developed a slice culture preparation from transgenic *Ascl1^{CreER/+};Rosa26^{ZsGreen/mTmG}* embryonic mouse lungs with NE progenitors labeled with ZsGreen and membrane GFP (mGFP) and other cells labeled with mTomato. This allowed tracking by time-lapse confocal microscopy of individual NE progenitors for 2–3 days in the 175-μm thick lung slices as they moved toward and integrated into target sites. This showed



progenitors migrating toward targets at $1\text{--}3 \mu\text{m/hr}$, dynamically extending processes and changing shape as they crawled over and around neighboring epithelial cells (Figures 6A–6M; Movie S1). Progenitors migrated tens of microns or more, traversing many neighboring epithelial cells or even crossing from one side of a developing bronchial tube to the other when the sides were apposed (Figures 6N–6V; Movie S2). Periodically, the migrating cells paused (“resting,” Figures 6I and 6M; Movie S1) or changed direction (Figure 6J, “extending away”; Movie S1), or displayed dynamic interactions with other NE cells (“kissing,” Figures 6F–6H; Movie S1) or the basement membrane (“dynamic anchors,” Figures 6R and 6S; Movie S2). After the progenitor approached the target, this slow and halting migration phase was followed by direct and rapid ($\sim 5 \mu\text{m/hr}$ for 2 hr) entry into the developing cluster. Live imaging confirmed this mode of NE cell migration predicted by the *in vivo* intermediates, and it revealed dynamic “exploratory” behaviors and interactions with

other NE cells and the basement membrane as migrating cells seek and ultimately converge at the target.

Transient Downregulation of Epithelial Adhesion and Polarity in Migrating Progenitors

The intermediates and movements imply that NE progenitors undergo changes in adhesion and polarity as they migrate and form a NEB. To begin to elucidate the underlying molecular events, we immunostained intermediates for adhesion and polarity markers (Figures 7A and 7C–7E). E-cadherin showed dynamic changes in expression during cluster formation (Figures 7A and 7C). The adhesion protein was expressed at normal levels in early progenitors but dramatically downregulated in migrating class 3 intermediates. Downregulation was transient because E-cadherin was again detected in class 5 intermediates that reached their target and re-established full contact with the basal lamina, although even then E-cadherin levels remained below that of

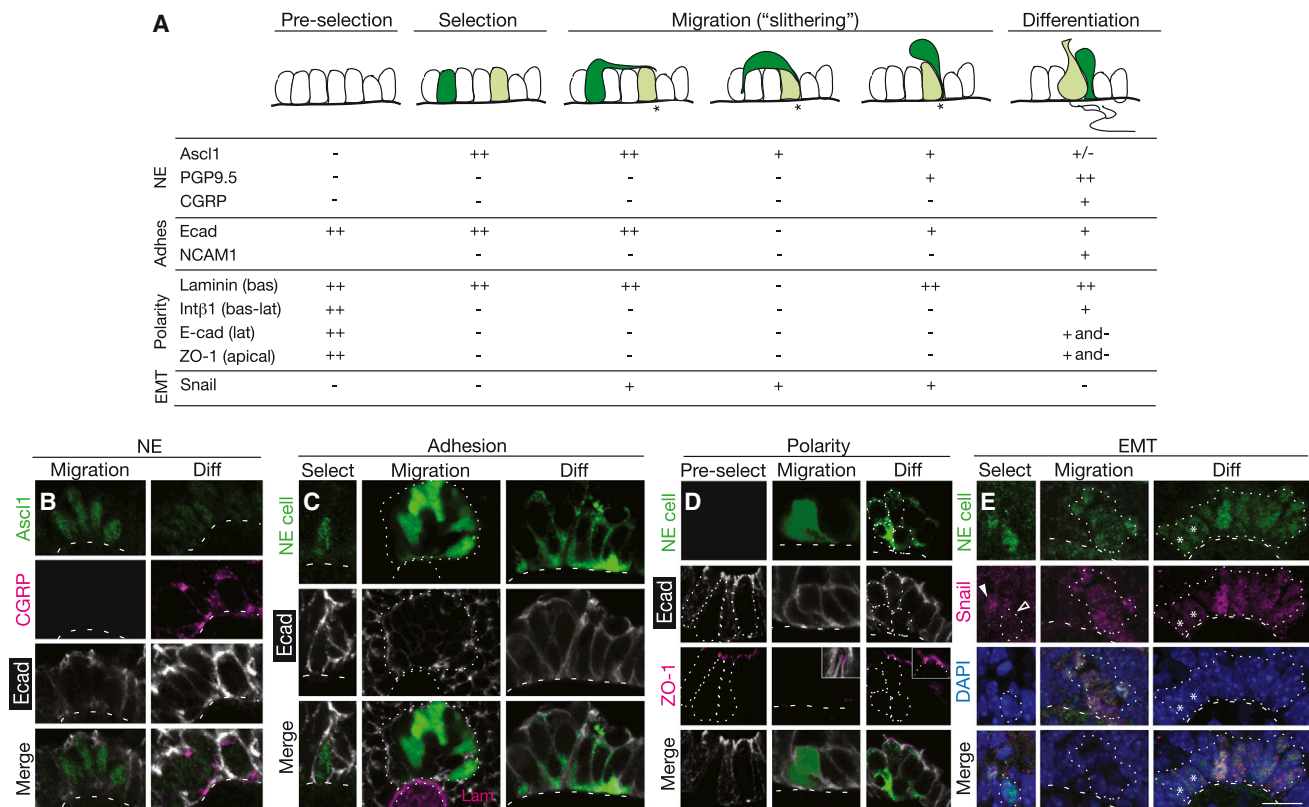


Figure 7. NE Progenitors Transiently Express EMT Regulator Snail and Downregulate Adhesion and Polarity Markers

(A) Expression of NE, adhesion (adhes), polarity, and EMT markers in NEB progenitors and intermediates (dark green) at steps indicated: before NE selection (pre-selection) and during NE selection (class 1), migration to form clusters (classes 2–4), and differentiation (class 5). Light green cells, maturing NE cell already at target (asterisk); white cells, NE cells before selection and other non-NE epithelial cells. Symbols show high (++), intermediate (+), low (+/-), no (-), or mixed (+ and -) expression. For polarity markers, expression refers to expression restricted to basal (bas), basolateral (bas-lat), or apical domains. See also Figure S7. (B–E) Confocal micrographs showing dynamic patterns of indicated NE (B), adhesion (C), polarity (D), and EMT (E) markers at indicated steps of NEB formation. Dots, outline of NEB; dashed line, basement membrane. (B) *Ascl1* (green) is expressed during migration (left panels; E13) but downregulated during differentiation (right panels; E18), whereas *CGRP* (magenta) is detected only during differentiation. (C) E-cadherin is expressed around entire plasma membrane during NE selection (*Ascl1*-positive (green) cell at E13; left panels), but is downregulated in class 3 and 4 intermediates (some of which are marked by *ZsGreen* in E15 *Ascl1*^{CreER/+}; *Rosa26*^{ZsGreen/+} mouse induced with tamoxifen at E12.5) in developing NEB (dotted outline; center panels). E-cadherin is again detected in NE cells (*CGRP*-positive; green) of mature NEB (PN 2 month; right panels), although expression is excluded from apical cell surface and detected along lateral boundaries in at least some cells. Cells neighboring NEB (not shown) express higher levels of E-cadherin and have larger apical surface that excludes E-cadherin. Lam, laminin γ 1 (magenta). (D) Before NE selection (E12; left panels) epithelial progenitors are uniformly polarized and tight junction protein ZO-1 (magenta) is detected at apical surface of each cell. ZO-1 is downregulated and undetected in intermediates (*Ascl1*^{CreER/+}; *Rosa26*^{ZsGreen/+} mouse labeled as above) during migration (center panels), when surrounding epithelial cells maintain its expression (inset), but is again detected during differentiation (right panels). (E) Snail proteins (magenta) are detected in some (filled arrowhead) but not other (open arrowhead) early NE progenitors (*Ascl1*-positive, green) during NE selection (E13; left panels), but all intermediates express Snail during migration (middle panels). Intermediates downregulate Snail in concert with *Ascl1* as clusters form and differentiate (E15; right panels), as seen by declining level of Snail and *Ascl1* in some cells (asterisks) within maturing cluster. Snail proteins are induced again later in NE cells (C.S.K., M.A.K., and J. Ouadah, unpublished data). Scale bar (B–F), 10 μ m.

surrounding non-NE epithelial cells. Immunostaining also showed induction of neural cell adhesion molecule (NCAM) during NEB formation. NCAM was expressed in class 5 intermediates (Figures 7A and S7B), implying a role in NEB assembly or cohesion.

Immunostaining for epithelial apicobasal polarity markers integrin β 1, laminin γ 1, and zona occludens-1 (ZO-1), and analysis of the apicobasal distribution of E-cadherin, showed that developing NE cells also undergo dynamic changes in polarity. Progenitors prior to NE cell selection (Figures 7D and S7A), and epithelial cells not selected (Figures S7C and S7D), show the expected api-

cobasal distribution of markers (Figure 7A). However, early NE progenitors (class 1 and 2) lose integrin β 1 and ZO-1 expression, and E-cadherin is distributed around the entire plasma membrane; the only apparent residual polarity is contact with the laminin γ 1-expressing basal lamina (Figure 7A). Migrating class 3 intermediates appear to have lost epithelial structure and polarity, with E-cadherin expression diminished as noted above and no contact with the basal lamina or only residual contact in local areas of low or undetectable laminin γ 1 (Figures 7C, 7D, and S7A). Class 4 intermediates begin to reestablish epithelial structure with E-cadherin again detected surrounding the entire cell

and their thin cytoplasmic extensions converging and contacting laminin in the basal lamina, suggesting re-initiation of apico-basal polarity (Figures 7A and 7D). Class 5 intermediates show continued maturation of epithelial structure and polarity that is complete in at least some cells, as evidenced by an expanded site of contact with laminin and basal lamina, expression of integrin $\beta 1$ (Figure S7A) along the entire basolateral surface, and expression of ZO-1 at the apical plasma membrane and exclusion of E-cadherin from this domain (Figures 7A and 7D).

These morphological and molecular changes show that NE progenitors undergo a transient epithelial-to-mesenchymal transition (EMT) during migration and NEB formation. Many EMT events are controlled by Snail proteins, zinc finger transcription factors that drive the switch (Mani et al., 2008; Strobl-Mazzulla and Bronner, 2012). Using an anti-Snail antiserum to detect Snail 1, 2, and 3, none was detected in the earliest NE progenitors or any other epithelial cells (Figures 7A and 7E). However, soon after initiation of *Ascl1* expression, nuclear Snail was detected and expression continued throughout cluster formation, before declining in class 5 intermediates as NE clusters form and re-establish epithelial structure and *Ascl1* levels decline. We conclude that early NE progenitors induce Snail expression and undergo EMT, although it is an unconventional EMT as progenitors remain associated with the epithelium and the process soon reverses as progenitors downregulate Snail and restore epithelial structure and polarity on reaching the target.

DISCUSSION

We have shown that NEBs form from NE progenitors initially distributed apparently randomly in the bronchial epithelium but then resolve in 2 days into clusters of 20–30 NE cells at stereotyped, diametrically opposed positions at the base of each bronchial branch. There is little or no progenitor proliferation during the process, so NEBs do not form by clonal expansion, as do *Drosophila* and other insect neurosensory organs. There is also little if any apoptosis of progenitors or reduction in progenitor number, so the clearing of progenitors from regions surrounding developing NEBs is not due to programmed cell death. Instead, lineage tracing, sparse labeling of progenitors, and high resolution imaging of intermediates in vivo, together with live imaging of progenitors in lung slices ex vivo, show that progenitors lose epithelial character and adopt fibroblast-like morphologies as they slowly (1–3 $\mu\text{m/hr}$ for 1–2 days) crawl over and around neighboring epithelial cells, converging at diametrically opposed sites at the base of each bronchial branch and then rapidly (1–2 hr) joining the cluster. The intermediates express EMT transcription factor Snail and transiently downregulate E-cadherin and epithelial polarity proteins and then re-express these adhesion and polarity proteins and turn on NE differentiation genes and NCAM as they re-establish epithelial structure and assemble into a neurosensory organ.

The results lead us to propose that NEBs form by a targeted mode of epithelial cell sorting we call “slithering,” in which the rearranging cells transiently lose epithelial structure and polarity yet remain intimately associated with the epithelial sheet as they traverse neighboring epithelial cells and converge at the target site. This mechanism of cell rearrangement differs dramati-

cally from intercalation, the classical mode of epithelial cell rearrangement, in which cells shrink lateral junctions with some neighboring cells while expanding junctions with others, exchanging positions while maintaining their polarized structure (Bertet et al., 2004; Blankenship et al., 2006; Guillot and Lecuit, 2013). At a mechanistic level, slithering is more reminiscent of cell delamination by EMT (Lamouille et al., 2014), such as during cancer metastasis (Scheel and Weinberg, 2012) or in neural crest formation where neural tube cells detach from the basement membrane and become mesenchyme-like cells that colonize distant sites (Strobl-Mazzulla and Bronner, 2012). During slithering, however, NE precursors never (or only transiently) leave the epithelial layer and they move comparatively short distances (tens of microns) and then rapidly regain epithelial character, so it is only a fleeting conversion. Epithelial cells can also undergo partial EMT during branching morphogenesis, in some organs forming highly dynamic, proliferative cell clusters at bud tips (Ewald et al., 2008), and individual proliferating epithelial cells can even occasionally send daughter cells to non-adjacent positions in the monolayer (“mitosis-associated cell dispersal”) (Packard et al., 2013). However, unlike these apparently random epithelial cell rearrangements and dispersals, slithering occurs without proliferation and is selective, directed, and purposeful. It results in active sorting of NE cells from surrounding epithelial cells in the monolayer, contrasting with the passive mechanisms long envisioned for cell sorting (Krens and Heisenberg, 2011; Steinberg, 1963) and forming neurosensory organs at defined positions.

Two prominent features of slithering are its selectivity for NE cells and their specific targeting to diametrically opposed positions at the base of each bronchial branch, raising the questions: what provides the guidance cue and its selectivity for NE cells? Because of the dynamic exploratory behavior of NE cells during slithering, we favor the idea that the guidance cue is a chemoattractant. Because only NE cells slither and they do so after turning on the master transcription factor *Ascl1*, we suggest that the chemoattractant receptor is specific for NE cells and part of their developmental program. However, there is no obvious cell or structure at target sites that could provide a point source of chemoattractant, and neurites that innervate NEBs arrive at their target only after the clusters form (Figures 2F and S3G–S3I). Perhaps the cue is a combination of more broadly distributed but overlapping signals, or even a physical signal related to the target structure: NEBs form at the most proximal and distal positions of the branch junction, where the connecting branches form the most obtuse (proximal) and acute (distal) angles. Although we have not systematically studied the formation of mini-clusters, available data indicate that they, too, form by slithering. However, because they do not form at stereotyped positions, their clustering could be driven by a homotypic attraction signal. A homotypic signal may also contribute to NEB formation, as real time imaging showed repeated transient contacts (“cell kisses”) among progenitors entering a NEB. A high priority now is to identify the signal(s) and their sources and receptors that control slithering and to determine if the same signals also guide outgrowth of neurites that target NEBs.

Because of the similarity in their developmental origins and clustered structures, slithering may also be used to form other

epithelial neurosensory organs such as taste buds (Chai et al., 2000; Okubo et al., 2009) and Merkel touch domes in skin that mediate light touch and spatial discrimination (Morrison et al., 2009; Van Keymeulen et al., 2009; Wright et al., 2015). Because slithering is difficult to detect without membrane marking or live imaging of individual intermediates, slithering could be a more widely used mechanism of movement and sorting of epithelial cells that has been overlooked even in well-studied epithelia such as the intestinal crypt, where Paneth cells move down from the proliferative zone as other cells move up (Battile et al., 2002). Slithering appears to be a late evolutionary innovation, at least in the lung, because although all mammals studied have NEBs, the most primitive extant fish species have only solitary pulmonary NE cells (Zaccone et al., 1989).

It will be important to identify the slithering program and indeed the full program of NEB formation including cell selection, migration, differentiation, and innervation, and how these processes go awry in Notch pathway mutants (Morimoto et al., 2012; Tsao et al., 2009) and other forms of pulmonary NE cell disease including those with excess or misplaced NE cells (Nassar et al., 2011; Young et al., 2011). One appealing idea is that the slithering program is transiently reactivated in their stem cell function as they move out of their niche to replace dying neighboring cells and permanently activated by mutation during oncogenesis. The latter could explain why small cell lung cancer arising from pulmonary NE cells metastasize early and are the most deadly form of lung cancer.

EXPERIMENTAL PROCEDURES

Animals

CD-1 was wild-type strain. Cre alleles and reporters are described in the [Supplemental Experimental Procedures](#). All animals were maintained and experiments performed in accordance with Stanford University's IACUC-approved protocols.

Immunohistochemistry and Histology

Embryos and lungs from timed pregnancies, with noon of the day of vaginal plug detection designated E0.5, were dissected, fixed in paraformaldehyde (PFA), and immunostains of whole mount lungs performed as described (Metzger et al., 2008) except as noted ([Supplemental Experimental Procedures](#)), then imaged by confocal microscopy or optical projection tomography. For cryosections, tissue was fixed in PFA or Zamboni's fixative, cryoprotected in 30% sucrose, frozen in OCT, and stored at -80°C . Frozen tissue blocks were sectioned and incubated sequentially with blocking solution, primary antibody, secondary antibody conjugated to Alexa fluorophores, and DAPI. For full description of antibodies and methods, see the [Supplemental Experimental Procedures](#).

Mapping Pulmonary NE Cells

Serial sections of entire left lobes ($n = 5$) of E16 CD-1 mice were stained for E-cadherin and Ascl1 and branches L.L1, L.L2, L.L3, L.L4, L.D1, L.D2, L.D3, and L.D4 systematically examined for NE cells. Additional ages and branch points in left and right lung lobes were analyzed as described ([Supplemental Experimental Procedures](#)).

Developmental Analysis of NE Cells

Serial sections of entire left lung lobes ($n = 5$ for each stage examined) of E12–E16 CD-1 mice were stained for Ascl1 and E-cadherin, and positions of NE cells were determined in branches and at junctions indicated. For NE cell counts ([Figure S3D](#)), serial sections of the entire left lobe were immunostained

for Ascl1 and E-cadherin and cells scored in each segment along left main bronchus as indicated.

For EdU incorporation, 300 μg EdU was injected intraperitoneally and lungs harvested 2 hr later. EdU was detected by click chemistry in cryosections immunostained for Ascl1 and counterstained with DAPI, then visualized by confocal microscopy. For phospho-histone H3 analysis, cryosections were immunostained for Ascl1 and anti-phospho-Histone 3-Ser10.

Lineage Tracing and Clonal Analysis

Lineage tracing was done with *Shh^{Cre/+};Rosa26^{mTmG/+}* mice (airway epithelial lineage) and with *Wnt1-Cre;Rosa26^{ZsGreen/+}* (neural crest lineage). Lungs were harvested between E17.5–E18.5, fixed, cryosectioned, and immunostained as above. To test monoclonality of NEBs ([Figure 3A](#)), cryosectioned E18 *Shh^{Cre/+};Rosa26^{Rbw/+}* lungs, in which all airway epithelial cells are permanently labeled early in development with one of the three Rainbow fluorescent reporters, were immunostained with anti-CGRP and an Alexa633-conjugated secondary. Cell number and colors in each NE cluster were assessed by confocal fluorescence microscopy. For clonal analysis, individual NE progenitors were sparsely labeled by tamoxifen injection at E11.5 (4 mg tamoxifen) or E12.5 (3 or 4 mg) of *Ascl1^{CreER/+};Rosa26^{Rbw/+}* mice, and lungs were harvested between E17.5–E18.5 and analyzed as above.

Labeling NE Developmental Intermediates

Ascl1^{CreER/+};Rosa26^{ZsGreen/mTmG} mice were induced with tamoxifen at E12.5, and lungs were harvested at E14.5 or E15.5 and serially sectioned, then co-stained for PGP9.5, E-cadherin and, for some samples, laminin gamma1 ($\gamma 1$). NEBs and mini-clusters were analyzed by confocal microscopy, and a 35–40 μm z stack was collected and each labeled NE progenitor examined at high resolution. Three-dimensional reconstructions (Velocity) of each progenitor were used to identify cell shape, cytoplasmic processes, and interactions with other cells and basement membrane. Detailed descriptions of the morphologic classes are given in the [Supplemental Experimental Procedures](#).

Live Imaging of NE Cells in Slice Culture

Ascl1^{CreER/+};Rosa26^{ZsGreen/mTmG} mice were induced with tamoxifen by oral gavage at E13 to label pulmonary NE progenitors with ZsGreen and mGFP and all other cells with TdTomato. At E15, lungs were harvested and left lobes and right caudal lobes separated, embedded in agarose and sectioned (Compressstone). Individual slices (175- μm thick) were transferred to a coverglass chamber, covered with Matrigel, and cultured in DMEM +F12 medium with 10% fetal bovine serum in an environmental chamber at 37°C and 5% CO_2 . After 2–3 hr to establish the culture, confocal images were collected every 20 min for 60 hr. See the [Supplemental Experimental Procedures](#).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, three tables, and two movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.021>.

AUTHOR CONTRIBUTIONS

C.S.K. and M.A.K. conceived project and designed experiments. C.S.K. executed and analyzed experiments. C.S.K. and M.A.K. interpreted data and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Makenna Morck for help with slice culture experiments, Jozeph Pendleton for help with mouse colony maintenance, Dr. Andrew Olson (Stanford Neuroscience Microscopy Service, NIH NS069375) for advice on video processing, Dr. Jane Johnson for *Ascl1^{CreER}* mice, members of the M.A.K. lab for discussions and comments on the manuscript, and Maria Petersen for help preparing the manuscript and figures. This work was supported by a grant from the NHLBI Progenitor Cell Biology Consortium (5U01HL099995), and a Tashia and John Morgridge Endowed Fellowship in Pediatric Translational

Medicine from the Stanford Child Health Research Institute (C.S.K.), Parker B. Francis Foundation Fellowship (C.S.K.), and NIH NHLBI K12 Scholar Award (C.S.K.). M.A.K. is an investigator of the Howard Hughes Medical Institute.

Received: April 13, 2015

Revised: July 8, 2015

Accepted: August 11, 2015

Published: October 1, 2015

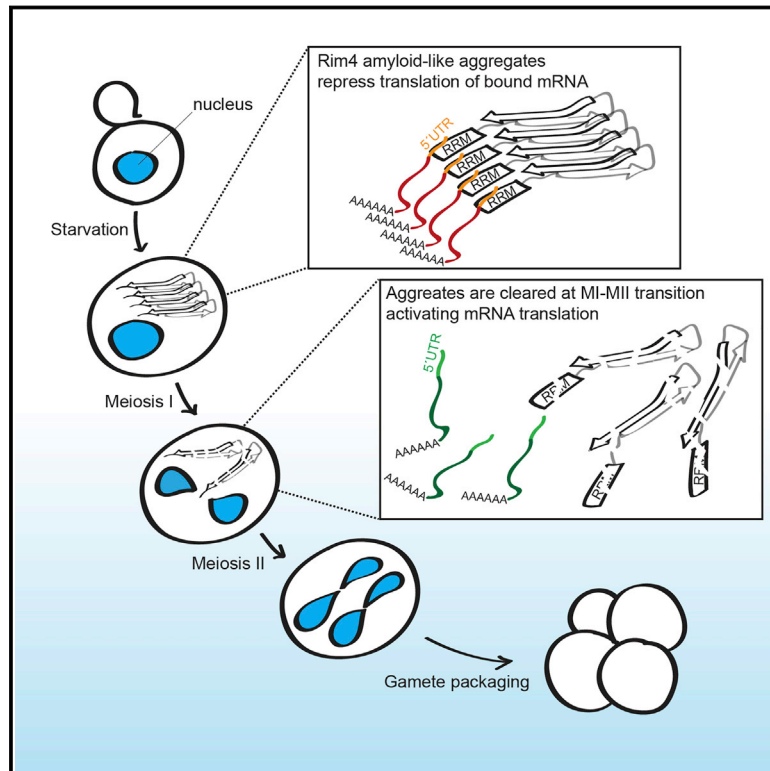
REFERENCES

- Aguiayo, S.M., Miller, Y.E., Waldron, J.A.J., Jr., Bogin, R.M., Sunday, M.E., Staton, G.W.J., Jr., Beam, W.R., and King, T.E.J., Jr. (1992). Brief report: idiopathic diffuse hyperplasia of pulmonary neuroendocrine cells and airways disease. *N. Engl. J. Med.* **327**, 1285–1288.
- Aigouy, B., Farhadifar, R., Staple, D.B., Sagner, A., Röper, J.-C., Jülicher, F., and Eaton, S. (2010). Cell flow reorients the axis of planar polarity in the wing epithelium of *Drosophila*. *Cell* **142**, 773–786.
- Anderson, D.J., and Axel, R. (1986). A bipotential neuroendocrine precursor whose choice of cell fate is determined by NGF and glucocorticoids. *Cell* **47**, 1079–1090.
- Battle, E., Henderson, J.T., Beghtel, H., van den Born, M.M.W., Sancho, E., Huls, G., Meeldijk, J., Robertson, J., van de Wetering, M., Pawson, T., and Clevers, H. (2002). Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. *Cell* **111**, 251–263.
- Bertet, C., Sulak, L., and Lecuit, T. (2004). Myosin-dependent junction remodeling controls planar cell intercalation and axis elongation. *Nature* **429**, 667–671.
- Blankenship, J.T., Backovic, S.T., Sanny, J.S.P., Weitz, O., and Zallen, J.A. (2006). Multicellular rosette formation links planar cell polarity to tissue morphogenesis. *Dev. Cell* **11**, 459–470.
- Brouns, I., Oztay, F., Pintelon, I., De Proost, I., Lembrechts, R., Timmermans, J.-P., and Adriaensen, D. (2009). Neurochemical pattern of the complex innervation of neuroepithelial bodies in mouse lungs. *Histochem. Cell Biol.* **131**, 55–74.
- Chai, Y., Jiang, X., Ito, Y., Bringas, P., Jr., Han, J., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127**, 1671–1679.
- Cubas, P., de Celis, J.F., Campuzano, S., and Modolell, J. (1991). Proneural clusters of achaete-scute expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes Dev.* **5**, 996–1008.
- Cutz, E., Perrin, D.G., Pan, J., Haas, E.A., and Krous, H.F. (2007). Pulmonary neuroendocrine cells and neuroepithelial bodies in sudden infant death syndrome: potential markers of airway chemoreceptor dysfunction. *Pediatr. Dev. Pathol.* **10**, 106–116.
- Cutz, E., Pan, J., Yeger, H., Domnik, N.J., and Fisher, J.T. (2013). Recent advances and controversies on the role of pulmonary neuroepithelial bodies as airway sensors. *Semin. Cell Dev. Biol.* **24**, 40–50.
- Deterding, R.R., Pye, C., Fan, L.L., and Langston, C. (2005). Persistent tachypnea of infancy is associated with neuroendocrine cell hyperplasia. *Pediatr. Pulmonol.* **40**, 157–165.
- Ewald, A.J., Brenot, A., Duong, M., Chan, B.S., and Werb, Z. (2008). Collective epithelial migration and cell rearrangements drive mammary branching morphogenesis. *Dev. Cell* **14**, 570–581.
- Feyrter, F. (1954). [Argyrophilia of bright cell system in bronchial tree in man]. *Z. Mikrosk. Anat. Forsch.* **61**, 73–81.
- Gillan, J.E., and Cutz, E. (1993). Abnormal pulmonary bombesin immunoreactive cells in Wilson-Mikity syndrome (pulmonary dysmaturity) and bronchopulmonary dysplasia. *Pediatr. Pathol.* **13**, 165–180.
- Gu, G., Dubauskaite, J., and Melton, D.A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447–2457.
- Guha, A., Vasconcelos, M., Cai, Y., Yoneda, M., Hinds, A., Qian, J., Li, G., Dickel, L., Johnson, J.E., Kimura, S., et al. (2012). Neuroepithelial body microenvironment is a niche for a distinct subset of Clara-like precursors in the developing airways. *Proc. Natl. Acad. Sci. USA* **109**, 12592–12597.
- Guillot, C., and Lecuit, T. (2013). Mechanics of epithelial tissue homeostasis and morphogenesis. *Science* **340**, 1185–1189.
- Harris, K.S., Zhang, Z., McManus, M.T., Harfe, B.D., and Sun, X. (2006). Dicer function is essential for lung epithelium morphogenesis. *Proc. Natl. Acad. Sci. USA* **103**, 2208–2213.
- Irvine, K.D., and Wieschaus, E. (1994). Cell intercalation during *Drosophila* germband extension and its regulation by pair-rule segmentation genes. *Development* **120**, 827–841.
- Ito, T., Udaka, N., Yazawa, T., Okudela, K., Hayashi, H., Sudo, T., Guillemot, F., Kageyama, R., and Kitamura, H. (2000). Basic helix-loop-helix transcription factors regulate the neuroendocrine differentiation of fetal mouse pulmonary epithelium. *Development* **127**, 3913–3921.
- Kim, E.J., Ables, J.L., Dickel, L.K., Eisch, A.J., and Johnson, J.E. (2011). *Ascl1* (*Mash1*) defines cells with long-term neurogenic potential in subgranular and subventricular zones in adult mouse brain. *PLoS ONE* **6**, e18472.
- Krens, S.F.G., and Heisenberg, C.-P. (2011). Cell sorting in development. *Curr. Top. Dev. Biol.* **95**, 189–213.
- Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196.
- Lauweryns, J.M., and Peuskens, J.C. (1972). Neuro-epithelial bodies (neuroreceptor or secretory organs?) in human infant bronchial and bronchiolar epithelium. *Anat. Rec.* **172**, 471–481.
- Lauweryns, J.M., and Van Lommel, A. (1987). Ultrastructure of nerve endings and synaptic junctions in rabbit intrapulmonary neuroepithelial bodies: a single and serial section analysis. *J. Anat.* **151**, 65–83.
- Lawrence, P.A. (1966). Development and determination of hairs and bristles in the milkweed bug, *Oncopeltus fasciatus* (Lygaeidae, Hemiptera). *J. Cell Sci.* **1**, 475–498.
- Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shiptsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715.
- Metzger, R.J., Klein, O.D., Martin, G.R., and Krasnow, M.A. (2008). The branching programme of mouse lung development. *Nature* **453**, 745–750.
- Morimoto, M., Nishinakamura, R., Saga, Y., and Kopan, R. (2012). Different assemblies of Notch receptors coordinate the distribution of the major bronchial Clara, ciliated and neuroendocrine cells. *Development* **139**, 4365–4373.
- Morrison, K.M., Miesegaes, G.R., Lumpkin, E.A., and Maricich, S.M. (2009). Mammalian Merkel cells are descended from the epidermal lineage. *Dev. Biol.* **336**, 76–83.
- Nassar, A.A., Jaroszewski, D.E., Helmers, R.A., Colby, T.V., Patel, B.M., and Mookadam, F. (2011). Diffuse idiopathic pulmonary neuroendocrine cell hyperplasia: a systematic overview. *Am. J. Respir. Crit. Care Med.* **184**, 8–16.
- Okubo, T., Clark, C., and Hogan, B.L.M. (2009). Cell lineage mapping of taste bud cells and keratinocytes in the mouse tongue and soft palate. *Stem Cells* **27**, 442–450.
- Packard, A., Georgas, K., Michos, O., Riccio, P., Cebrian, C., Combes, A.N., Ju, A., Ferrer-Vaquer, A., Hadjantonakis, A.-K., Zong, H., et al. (2013). Luminal mitosis drives epithelial cell dispersal within the branching ureteric bud. *Dev. Cell* **27**, 319–330.
- Park, K.-S., Liang, M.-C., Raiser, D.M., Zamponi, R., Roach, R.R., Curtis, S.J., Walton, Z., Schaffer, B.E., Roake, C.M., Zmoos, A.-F., et al. (2011). Characterization of the cell of origin for small cell lung cancer. *Cell Cycle* **10**, 2806–2815.
- Polak, J.M., Becker, K.L., Cutz, E., Gail, D.B., Goniakowska-Witalinska, L., Gosney, J.R., Lauweryns, J.M., Linnoila, I., McDowell, E.M., Miller, Y.E., et al. (1993). Lung endocrine cell markers, peptides, and amines. *Anat. Rec.* **236**, 169–171.

- Reynolds, S.D., Hong, K.U., Giangreco, A., Mango, G.W., Guron, C., Morimoto, Y., and Stripp, B.R. (2000). Conditional clara cell ablation reveals a self-renewing progenitor function of pulmonary neuroendocrine cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **278**, L1256–L1263.
- Rinkevich, Y., Lindau, P., Ueno, H., Longaker, M.T., and Weissman, I.L. (2011). Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. *Nature* **476**, 409–413.
- Rock, J.R., and Hogan, B.L.M. (2011). Epithelial progenitor cells in lung development, maintenance, repair, and disease. *Annu. Rev. Cell Dev. Biol.* **27**, 493–512.
- Scheel, C., and Weinberg, R.A. (2012). Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Semin. Cancer Biol.* **22**, 396–403.
- Scheuermann, D.W. (1987). Morphology and cytochemistry of the endocrine epithelial system in the lung. *Int. Rev. Cytol.* **106**, 35–88.
- Song, H., Yao, E., Lin, C., Gacayan, R., Chen, M.-H., and Chuang, P.-T. (2012). Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis. *Proc. Natl. Acad. Sci. USA* **109**, 17531–17536.
- Steinberg, M.S. (1963). Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science* **141**, 401–408.
- Strobl-Mazzulla, P.H., and Bronner, M.E. (2012). Epithelial to mesenchymal transition: new and old insights from the classical neural crest model. *Semin. Cancer Biol.* **22**, 411–416.
- Sutherland, K.D., Proost, N., Brouns, I., Adriaensen, D., Song, J.-Y., and Berns, A. (2011). Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. *Cancer Cell* **19**, 754–764.
- Tsao, P.-N., Vasconcelos, M., Izvolsky, K.I., Qian, J., Lu, J., and Cardoso, W.V. (2009). Notch signaling controls the balance of ciliated and secretory cell fates in developing airways. *Development* **136**, 2297–2307.
- Van Keymeulen, A., Mascré, G., Youseff, K.K., Harel, I., Michaux, C., De Geest, N., Szpalski, C., Achouri, Y., Bloch, W., Hassan, B.A., and Blanpain, C. (2009). Epidermal progenitors give rise to Merkel cells during embryonic development and adult homeostasis. *J. Cell Biol.* **187**, 91–100.
- van Meerbeeck, J.P., Fennell, D.A., and De Ruyscher, D.K.M. (2011). Small-cell lung cancer. *Lancet* **378**, 1741–1755.
- Wright, M.C., Reed-Geaghan, E.G., Bolock, A.M., Fujiyama, T., Hoshino, M., and Maricich, S.M. (2015). Unipotent, Atoh1+ progenitors maintain the Merkel cell population in embryonic and adult mice. *J. Cell Biol.* **208**, 367–379.
- Young, L.R., Brody, A.S., Inge, T.H., Acton, J.D., Bokulic, R.E., Langston, C., and Deutsch, G.H. (2011). Neuroendocrine cell distribution and frequency distinguish neuroendocrine cell hyperplasia of infancy from other pulmonary disorders. *Chest* **139**, 1060–1071.
- Zaccone, G., Tagliaferro, G., Goniakowska-Witalinska, L., Fasulo, S., Ainis, L., and Mauceri, A. (1989). Serotonin-like immunoreactive cells in the pulmonary epithelium of ancient fish species. *Histochemistry* **92**, 61–63.

Regulated Formation of an Amyloid-like Translational Repressor Governs Gametogenesis

Graphical Abstract



Authors

Luke E. Berchowitz, Greg Kabachinski, Margaret R. Walker, ..., Wendy V. Gilbert, Thomas U. Schwartz, Angelika Amon

Correspondence

angelika@mit.edu

In Brief

Amyloid-like aggregation of the RNA-binding protein Rim4 controls gametogenesis by repressing message-specific translation.

Highlights

- Amyloid-like aggregation of the RNA-binding protein Rim4 controls gametogenesis
- The amyloid-like form of Rim4 is active and represses translation
- Aggregation and clearance of Rim4 are developmentally regulated
- Amyloid-like aggregation of RNA-binding proteins during gametogenesis is conserved

Regulated Formation of an Amyloid-like Translational Repressor Governs Gametogenesis

Luke E. Berchowitz,¹ Greg Kabachinski,² Margaret R. Walker,¹ Thomas M. Carlile,² Wendy V. Gilbert,² Thomas U. Schwartz,² and Angelika Amon^{1,2,*}

¹David H. Koch Institute for Integrative Cancer Research and Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

*Correspondence: angelika@mit.edu

<http://dx.doi.org/10.1016/j.cell.2015.08.060>

SUMMARY

Message-specific translational control is required for gametogenesis. In yeast, the RNA-binding protein Rim4 mediates translational repression of numerous mRNAs, including the B-type cyclin *CLB3*, which is essential for establishing the meiotic chromosome segregation pattern. Here, we show that Rim4 forms amyloid-like aggregates and that it is the amyloid-like form of Rim4 that is the active, translationally repressive form of the protein. Our data further show that Rim4 aggregation is a developmentally regulated process. Starvation induces the conversion of monomeric Rim4 into amyloid-like aggregates, thereby activating the protein to bring about repression of translation. At the onset of meiosis II, Rim4 aggregates are abruptly degraded allowing translation to commence. Although amyloids are best known for their role in the etiology of diseases such as Alzheimer's, Parkinson's, and diabetes by forming toxic protein aggregates, our findings show that cells can utilize amyloid-like protein aggregates to function as central regulators of gametogenesis.

INTRODUCTION

Production of haploid gametes from diploid progenitor cells is a hallmark of sexual reproduction. The process is mediated by a specialized cell division, meiosis, where two chromosome divisions, meiosis I and II, follow a single S phase. While meiosis II resembles mitosis in that sister chromatids are separated, meiosis I is unusual in that homologous chromosomes rather than sister chromatids are segregated. Message-specific translational control of cyclins is required for progression through the meiotic divisions (reviewed in [Kronja and Orr-Weaver, 2011](#)). In budding yeast, translational control of the B-type cyclin *CLB3* is essential for establishing the unusual meiosis I division. *CLB3* mRNA is translationally repressed during meiosis I but is translated during meiosis II ([Carlile and Amon, 2008](#)). When *CLB3* is inappropriately expressed during meiosis I, meiosis I chromosome segregation is suppressed. Instead, cells undergo

a mitosis-like division illustrating the importance of translational control for gamete formation ([Miller et al., 2012](#)).

During meiosis I, the RNA-binding protein Rim4 represses the translation of *CLB3* and numerous other mRNAs, the protein products of which are required during the late stages of gametogenesis ([Berchowitz et al., 2013](#)). Rim4 binds to the 5'UTR of *CLB3* mRNA and represses translation until the onset of meiosis II. At this stage of meiosis, Rim4 is abruptly degraded. This degradation is controlled by the meiosis-specific protein kinase Ime2 ([Berchowitz et al., 2013](#)). These findings made it clear that understanding how Rim4 represses translation is essential to understanding how meiosis and gametogenesis are governed. Here, we shed light on this question. We show that Rim4 forms aggregates to repress translation. Rim4 aggregates are SDS-resistant, bind biotinylated isoxazole (b-isox), and can be pelleted and sedimented from detergent-containing lysate by ultracentrifugation. In vitro, Rim4 forms fibrillar structures containing β sheet. These biochemical properties are indicative of fiber-forming leading us to define Rim4 as "amyloid-like."

Amyloids and amyloid-like aggregates have been implicated in processes such as hormone storage ([Maji et al., 2009](#)) and melanin production in mammals ([Fowler et al., 2006](#)) and heterokaryon formation in filamentous fungi ([Dos Reis et al., 2002](#)). In *Drosophila*, the Orb2 protein forms oligomers with amyloid-like properties that play a critical role in long-term memory persistence ([Majumdar et al., 2012](#)). In mammals, RIP1/RIP3 protein kinases form a heterotypic amyloid complex that is critical for programmed cell necrosis ([Li et al., 2012](#)).

While some biological functions have been ascribed to amyloids, these aggregates are predominantly understood in the context of human diseases such as Alzheimer's, Parkinson's, amyotrophic lateral sclerosis (ALS), diabetes, and all known prion diseases (reviewed in [Knowles et al., 2014](#)). Since their discovery over 150 years ago ([Sipe and Cohen, 2000](#)), the main focus of amyloid research has been on understanding the proteins' disease-causing properties. In most, if not all, diseases associated with amyloids, the aggregates have been proposed to function in a dominant negative manner, sequestering monomeric species into the aggregates causing their depletion within cells. In addition to depleting functional proteins, amyloids are thought to interfere with cellular functions ([Thomas et al., 1996](#)), cause cell death ([Loo et al., 1993](#)), and in cases where they are secreted, induce inflammation ([Akiyama et al., 2000](#)).

Additionally, accumulating evidence suggests that soluble oligomeric amyloid precursors are critical drivers of neurotoxicity and are often described as inherently toxic (Bucciantini et al., 2002; Walsh and Selkoe, 2007). Here, we show that proteins with amyloid-like properties can control gene expression. We demonstrate that the translational repressor Rim4 forms large assemblies with amyloid-like biochemical properties and that it is the aggregated form of Rim4 that is the active, translationally repressive form of the protein. Remarkably, assembly of these structures is developmentally regulated. Upon initiation of sporulation, one of the inducers of the meiotic fate, starvation, induces the switch from monomeric Rim4 into amyloid-like Rim4 rendering the protein competent to repress *CLB3* translation. At the onset of meiosis II, translational repression of *CLB3* is lifted through the degradation of Rim4 aggregates. Our results demonstrate that regulated formation of amyloid-like aggregates establishes and upholds gametogenesis.

RESULTS

Rim4 Is Efficiently Precipitated from Meiotic Lysate by B-Isox

To begin to understand how Rim4 represses translation we used computational approaches to identify potential functional domains within the protein. Rim4 contains three N-terminal RNA-recognition motifs (RRMs) that are required for RNA binding and translational repression (Berchowitz et al., 2013). The protein also contains a C-terminal low-complexity (LC) region (Figure 1A) and harbors two poly-N stretches and a computationally predicted prion domain (Alberti et al., 2009; Soushko and Mitchell, 2000). Several RNA-binding proteins containing LC sequences have the ability to form amyloid-like fibrous aggregates. Additionally, LC sequences/prion domains are important for these proteins to be recruited into nonmembrane-bound intracellular compartments such as stress granules (Gilks et al., 2004) that have been shown to function in translational control (Kedersha et al., 1999).

Many fiber-forming RNA-binding proteins can be precipitated from aqueous solution or cell extracts by binding to the compound b-isox via their LC sequences (Kato et al., 2012). To determine whether Rim4 exhibits this property and thus shares biochemical similarity with other fiber-forming RNA-binding proteins, we asked whether Rim4 binds to b-isox. In order to obtain a uniform population of cells in which Rim4 inhibits translation, we induced cells lacking the transcription factor Ndt80 to undergo sporulation. Cells lacking Ndt80 progress through the early stages of the meiotic cell cycle efficiently but arrest in G2 prior to meiosis I entry (Xu et al., 1995). In this cell-cycle stage, Rim4 is active and translation of *CLB3* is inhibited (Berchowitz et al., 2013). We found that b-isox efficiently precipitated Rim4 from lysates prepared from cells lacking Ndt80 (Figure 1B).

To determine more precisely when during meiosis b-isox was able to precipitate Rim4, we examined cells induced to progress through a synchronous meiosis using the *NDT80* block-release system. In this system, *NDT80* expression is driven from the *GAL1-10* promoter (*pGAL-NDT80*) that in turn is controlled by an estradiol-regulatable Gal4-ER fusion (*GAL4.ER*). *pGAL-NDT80* cells will arrest in meiotic G2 in sporulation medium

lacking estradiol but will progress synchronously through the meiotic divisions upon estradiol addition (Benjamin et al., 2003; Carlile and Amon, 2008). Rim4 was efficiently precipitated by b-isox for 2 hr following the release from the G2 block, when cells were in meiosis I (Figures 1C and 1D). As shown previously (Berchowitz et al., 2013), when cells enter meiosis II, Rim4 was degraded (Figures 1C and 1D).

To assess the specificity of b-isox, we used quantitative mass spectrometry to query the yeast proteome for meiotic b-isox binding proteins. We released cells from the *NDT80* block and collected samples spanning the two meiotic divisions (Figure S1). B-isox was remarkably selective. Rim4 was the most enriched protein during early meiosis I (1 hr after release from the G2 block; Figures 1E and 1F) and only three other proteins, among them the prion Sup35, exhibited significant enrichment at this time (Figure 1E; Table S1). We conclude that Rim4, like prions and other proteins with amyloid-like properties, binds to b-isox and does so during the stages of meiosis when the protein actively inhibits *CLB3* translation.

Rim4 Forms Amyloid-like Aggregates In Vitro

Yeast prions and other amyloidogenic proteins such as Huntingtin form amyloid aggregates via regions within the protein rich in N and Q residues (DePace et al., 1998; Scherzinger et al., 1997). Given that Sup35 and Rim4 were efficiently precipitated by b-isox and that Rim4 contains poly-N stretches, we next asked whether Rim4, like Sup35, forms amyloid-like aggregates in vitro. We generated highly purified recombinant Rim4 harboring minimal nucleic acid contamination (Figures S2A–S2C) and analyzed concentrated protein by transmission electron microscopy (TEM). At a concentration of 40 mg/ml, Rim4 formed fibrillar helical structures with a fixed width of ~15 nm and length up to 15 μ m visible by both negative stain and cryoelectron TEM (cryo-TEM) (Figure 2A). The ability to form fibers depended on the protein's C-terminal LC sequences. A mutant derivative lacking the C-terminal 294 amino acids but leaving the three RRM intact (Rim4 Δ 294C) did not form long helical fibers (Figure 2B). Because Rim4 is an RNA-binding protein (Berchowitz et al., 2013), we examined the effects of RNA on Rim4 fiber formation. Incubation of non-fibrous Rim4 with a bona fide RNA target, the *CLB3* 5'UTR, or total RNA did not induce fiber formation as assessed by TEM (Figure S2D). Furthermore, RNase treatment of pre-formed fibers did not result in fiber disassembly (Figure S2E). These results indicate that fiber formation in vitro can occur in the absence of RNA.

Biochemical analyses further demonstrated that Rim4 fibers are composed of β sheets, a hallmark of amyloid. Circular dichroism (CD) spectrum analysis of fibrous Rim4 revealed a minimum at 218 nm (Figure S2F), which is characteristic of structures rich in β sheet. In contrast, the CD spectrum of Rim4 Δ 294C was predominated by α -helical signature (Figure S2F). Furthermore, Rim4 but not Histone H1 bound to Thioflavin T (ThT), a dye that stains the β sheet-rich structures of amyloid and not monomers (Nilsson, 2004) (Figure 2C). Rim4 lacking the C-terminal LC sequences also bound Thioflavin T but with reduced affinity compared to full-length protein (Figure 2C). We speculate that the N-terminal disordered sequence (Figure 1A) mediates the production of some β sheet structures in the

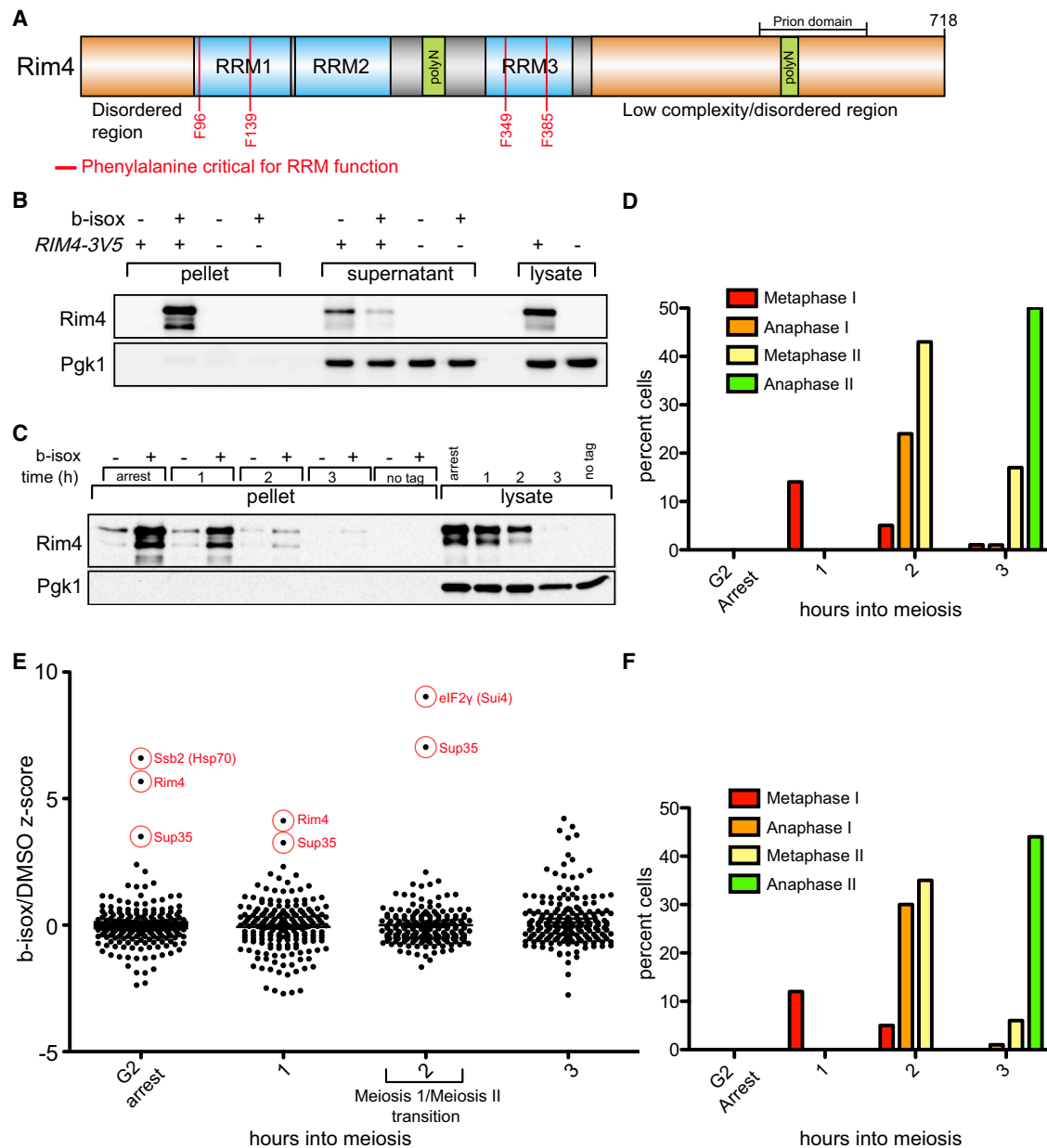


Figure 1. B-Isox Binds Rim4 in Meiosis I

(A) Diagram of Rim4.

(B) *RIM4-3V5*, *pGAL-NDT80*, and *GAL4.ER* (A30868) cells were induced to sporulate at 30°C. After 6 hr when cells had arrested in G2 due to the lack of Ndt80, lysates were prepared. Lysates were incubated with 100 μ M b-isox (or DMSO control) and Rim4 abundance was examined by western blot analysis. Pgk1 did not bind b-isox.

(C) Cells were grown as in (B) except cells were released from the G2 block by the addition of 1 μ M β -estradiol. Lysates were prepared from G2-arrested cells (0 hr) and at 1, 2, and 3 hr post-release and analyzed as in (B).

(D) Meiotic staging data for (C). The percentage of metaphase I, anaphase I, metaphase II, and anaphase II cells was determined.

(E) *pGAL-NDT80* (A15055) strains carrying the *GAL4.ER* were grown and lysates prepared as in (C). Precipitated proteins were ITRAQ-labeled and analyzed by mass spectrometry. Shown are the ratios of b-isox enrichment/DMSO for each identified protein converted to Z scores.

(F) Meiotic staging data from (E).

See also Figure S1 for detailed workflow and Tables S1 and S2.

truncated protein. In concordance with these results we observed that Rim4 is recognized by an α -amyloid fibril antibody (α -amyloid fibrils OC, Millipore) (Figure 2D). Rim4 Δ 294C is

recognized to a much lesser extent and Histone H1 not at all. We note that unlike other LC-sequence containing proteins (Kato et al., 2012), Rim4 did not form hydrogels, even when

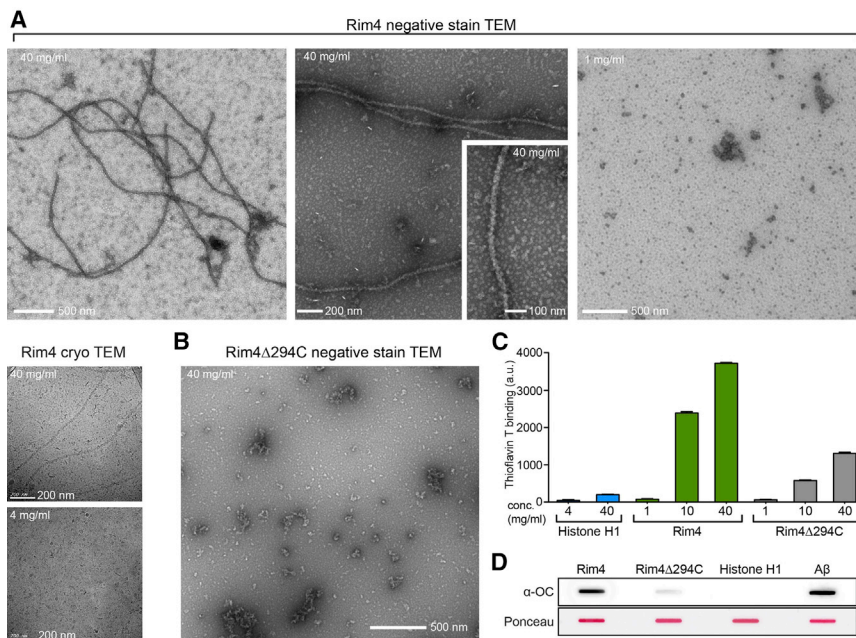


Figure 2. Rim4 Forms Amyloid-like Aggregates In Vitro

(A and B) Transmission electron microscopy (TEM) of fibers formed by recombinant Rim4 and Rim4Δ294C (aa 1–420) incubated at 4°C overnight. The top three panels are micrographs depicting negatively stained samples (prepared at 40 mg/ml and 1 mg/ml). The bottom left panels depict Rim4 samples analyzed by cryo-EM at concentrations of 40 mg/ml (left) and 4 mg/ml (right). Scale bars are shown in white.

(C) Recombinant Rim4 and Rim4Δ294C (various concentrations diluted to 1 mg/ml just before assay) were incubated with 16 ng/ml Thioflavin T, and the fluorescence intensity was measured as in Nilsson (2004). Histone H1 at 4 and 40 mg/ml was used as a control.

(D) Recombinant protein (10 μg) was slot-blotted onto nitrocellulose and assayed for reactivity with α-amyloid fibril OC. Rim4, Rim4Δ294C, and Histone H1 concentrated to 40 mg/ml were assayed. Amyloid beta (Aβ, Anaspec) concentrated to 0.25 mg/ml was used as a positive control.

See also Figure S2 and Table S2.

incubated at 4°C for months at high concentrations. Nevertheless, our analyses demonstrate that recombinant Rim4 forms aggregates with several morphological and biochemical properties of amyloid in the absence of other proteins or nucleic acids and that the C-terminal LC sequences are critical for fiber formation.

Evidence that Rim4 Forms Aggregates In Vivo

Amyloid-like aggregates can form stable structures in vivo as judged by fluorescence recovery after photobleaching (FRAP) (Kayatekin et al., 2014). For example, when the amyloidogenic Huntingtin protein harboring a greatly expanded poly-Q stretch (Htt103Q) is expressed in yeast, aggregated foci display very little signal recovery after photobleaching whereas non-aggregating Huntingtin proteins exhibit rapid signal recovery (Kayatekin et al., 2014). To determine whether Rim4 forms aggregates in vivo, we analyzed the FRAP kinetics of wild-type and aggregate-defective Rim4 that lacks the C-terminal 204 amino acids (Rim4Δ204C) in sporulating cells (Figures 3A and 3B). After photobleaching, wild-type Rim4 showed minimal recovery while the Rim4Δ204C signal recovered with similar kinetics as the soluble cytoplasmic enzyme Pgc1. This result demonstrates that Rim4 structures within the cell are largely static. This rigidity depends on Rim4's C-terminal 204 amino acids.

The localization pattern of Rim4 is also consistent with the protein forming aggregates in cells during stages of meiosis when the protein inhibits translation. In meiotic G2-arrested cells, Rim4 protein levels were similar between wild-type and Rim4Δ204C as judged by signal intensity within a fixed area of the cell (Figures 3C and 3D). However, whereas the Rim4Δ204C signal was uniformly distributed throughout the cytoplasm, wild-type Rim4 appeared more focal, leading to a more granular staining pattern (Figure 3C). This was most evident when we analyzed heterogeneity in pixel intensity by comparing the SDs of the mean pixel signal intensity within a defined area

of cytoplasm. The SD was significantly higher for wild-type Rim4 than the aggregate-defective Rim4Δ204C mutant protein (Figure 3E).

To determine whether the C-terminal LC sequences of Rim4 affect Rim4 clearance, we compared protein levels of wild-type Rim4 and the Rim4Δ204C mutant in live cells undergoing meiosis. To ensure that the Rim4 mutant protein did not affect meiotic progression, we conducted the experiment in diploid cells that carried one wild-type copy of *RIM4* and either one EGFP-tagged wild-type *RIM4* or *rim4Δ204C* allele. Rim4 accumulated during premeiotic S phase and persisted throughout meiosis I at which point Rim4 was abruptly degraded (Figures 3F, S3A, and S3B; Movies S1, S2, S3, and S4). Surprisingly, Rim4Δ204C was not cleared from cells efficiently compared to Rim4-EGFP (Figures 3F and S3C; Movies S5 and S6). The mutant protein persisted in cells well beyond exit from meiosis I. This result indicates that the C-terminal 204 amino acids are important for Rim4 aggregation in vivo and contain elements necessary for proper Rim4 clearance at the onset of meiosis II.

Rim4 Forms SDS-Resistant Aggregates during Meiosis

To further examine whether Rim4 indeed forms amyloid-like aggregates in vivo, we employed semi-denaturing detergent agarose gel electrophoresis (SDD-AGE). Amyloid (as opposed to amorphous or globular) aggregates are SDS-resistant and SDD-AGE allows for the resolution of these structures (Alberti et al., 2009). Rim4 formed SDS-resistant aggregates in cells progressing through meiosis in a synchronous manner (Figure 4A). Previously described fiber-forming (hydrogelling) RNA-binding proteins do not form SDS-resistant structures (Kato et al., 2012). Thus, Rim4 forms different structures from the fiberizing RNA-binding proteins that form hydrogels. However, we cannot rule out that Rim4 aggregates adopt a higher-order static hydrogel state within the cell.

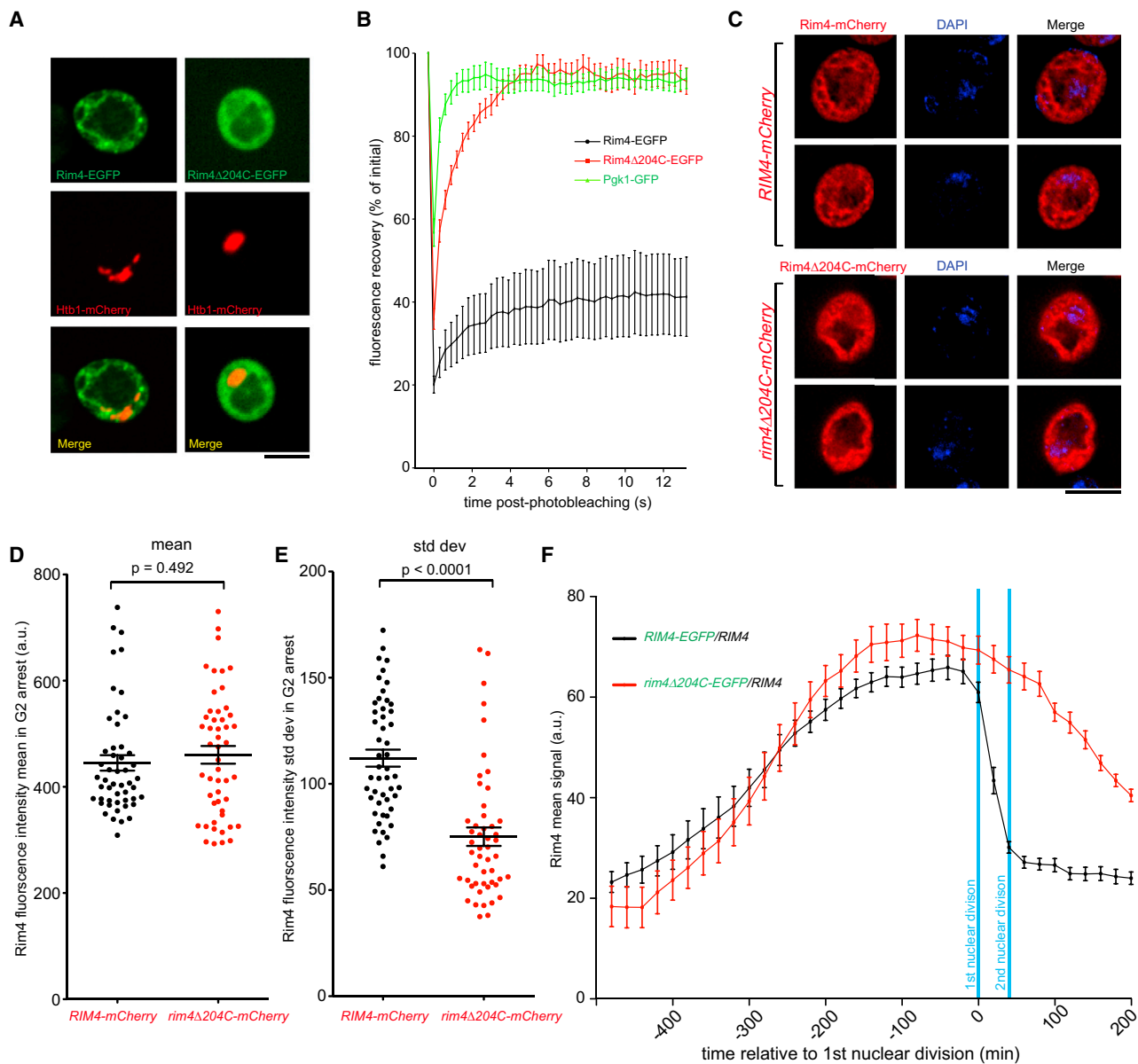


Figure 3. Evidence that Rim4 Forms Aggregates in Cells

(A and B) Rim4 forms static aggregates in cells. Strains heterozygous for *HTB1-mCherry* and homozygous for either *RIM4-EGFP* (A36245) or *rim4Δ204C-EGFP* (A36706) were induced to sporulate at 30°C. The mobility of Rim4 aggregates was determined by FRAP of live cells. Cells were imaged starting at 5 hr after transfer into sporulation medium. A haploid mitotic culture expressing *Pgk1-GFP* (A34422) was used as a soluble control. Bleached areas were normalized to an unbleached control area. Time point zero was the first observed time point after bleaching. Representative cells are shown in (A) with Rim4 in green and the nucleus in red (scale bar, 5 μm). Signal recovery over time is shown in (B). Plotted is the mean and SEM of ten cells per strain.

(C–E) *pGAL-NDT80*, *GAL4.ER* cells harboring *RIM4-mCherry* (A35632), or *rim4Δ204C-mCherry* (A37237) were induced to sporulate at 30°C. After 6 hr when cells had arrested in G2 due to the lack of *Ndt80*, the cells were fixed and imaged. (C) Single-plane images of representative cells arrested in G2 are shown. Scale bar, 5 μm. (D and E) Rim4-mCherry signal (vacuole excluded) was quantified in the central z plane. The mean Rim4 signal is not significantly different between the two strains, however, the SD (aggregation proxy) is significantly different. *n* = 50 cells per strain.

(F) Cells either harboring a *RIM4-EGFP* (A36775) or *rim4Δ204C-EGFP* (A36773) fusion were induced to sporulate at 30°C. After 1 hr of growth in batch culture, the cells were loaded onto a microfluidics chip and imaged every 20 min. A quantification of Rim4-EGFP signal over time is shown. Plots were aligned according to the first meiotic division. Cells harboring *RIM4-EGFP* and *rim4Δ204C-EGFP* are shown in black and red, respectively. Plotted is the mean and SEM of ten cells per strain. Scale bar, 5 μm.

See also [Figure S3](#), [Table S2](#), and [Movies S1](#), [S2](#), [S3](#), [S4](#), [S5](#), and [S6](#).

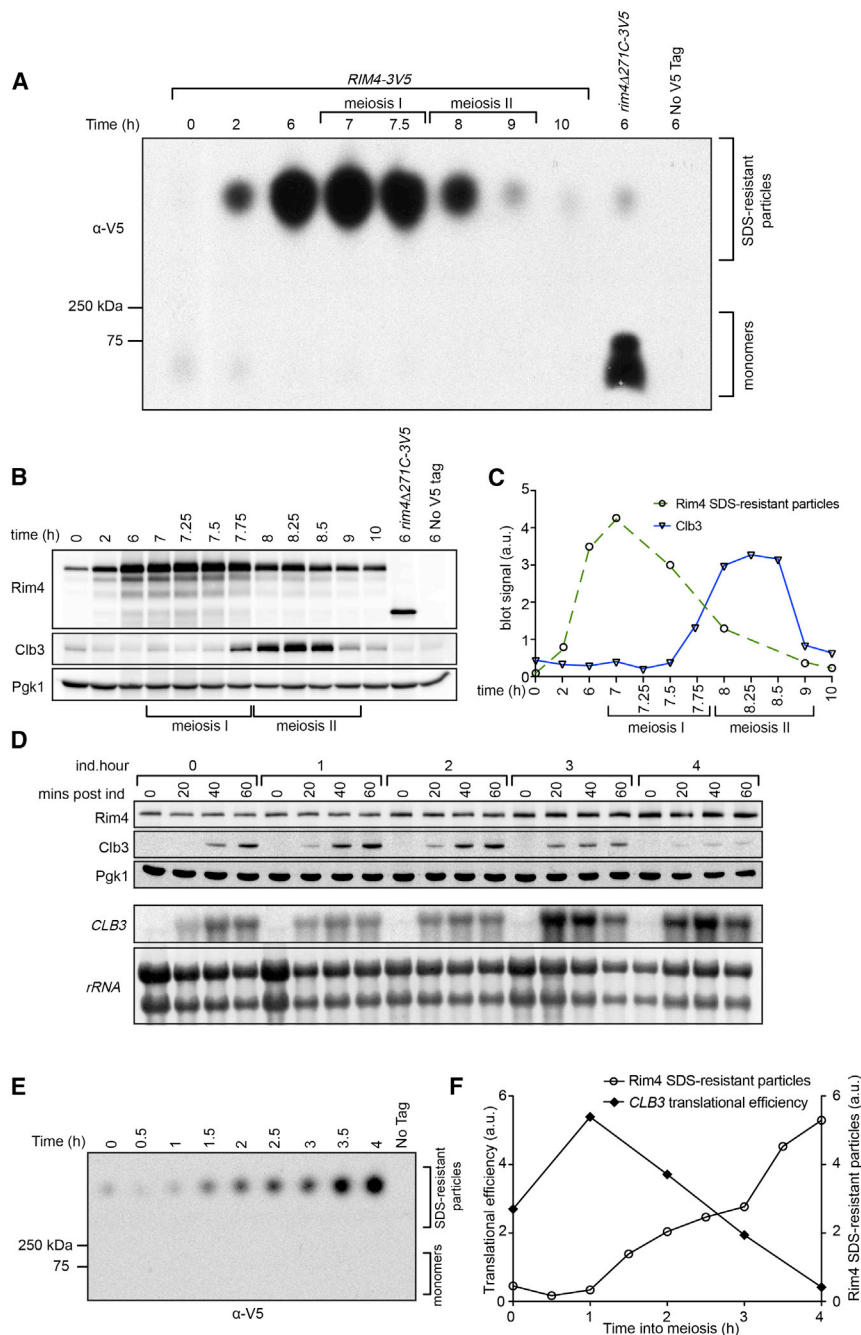


Figure 4. Rim4 Forms Amyloid-like Aggregates In Vivo to Repress Translation

(A) Rim4 amyloid-like aggregates form during premeiotic S phase that are cleared during meiosis II. *pGAL-NDT80* cells (A30868) were induced to sporulate and Rim4's ability to form SDS-resistant aggregates was analyzed as in Alberti et al. (2009). Samples were incubated in 2% SDS for 10 min and resolved on 1.7% agarose (with 0.1% SDS), and Rim4 was detected by immunoblot analysis. A strain harboring Rim4 lacking its C-terminal LC sequences (A33848) is shown as a control.

(B and C) Western blot analysis of time course shown in (A). Total levels of Rim4-3V5, Clb3-3HA, and Pgk1 (loading control) are shown in (B). Quantification of blots is shown in (C). Rim4 SDS-resistant aggregates were quantified from SDD-AGE in (A) and Clb3 from SDS-PAGE in (B). Total levels of Rim4 and Clb3 were normalized to Pgk1. (D–F) *CLB3* translational repression begins between 2 and 3 hr after induction of sporulation (see Figure S4A for detailed workflow). A culture harboring *RIM-3V5*, *GAL4.ER*, and *pGAL-5'UTR_{CLB3}-CLB3* (A35430) was split and induced to sporulate at 30°C. *pGAL-5'UTR_{CLB3}-CLB3* was induced by β-estradiol at the indicated time. Rim4-3V5, Clb3-3HA, and Pgk1 (loading control) protein and *CLB3* mRNA and *rRNA* (loading control) levels were determined at the indicated times following β-estradiol addition (D). SDS-resistant Rim4 aggregates were analyzed in (E). A quantification of Rim4 aggregate formation (open circles) and Clb3 translation (Clb3 protein/*CLB3* mRNA after 60 min induction: closed circles) is shown in (F). See also Table S2.

could originate from cells that either did not enter sporulation or arrest at some point in meiosis prior to Rim4 aggregate formation. Second, Rim4 aggregates could be disassembled in meiosis II before they are degraded. We note that monomeric Rim4 is not detected in SDD-AGE in these late meiotic time points, but previous studies indicated that SDD-AGE does not reliably detect monomeric proteins (Halfmann et al., 2012). It is also possible that monomeric Rim4 is unstable (indeed the protein is degraded during meiosis II) and hence

degraded during the lengthy SDD-AGE extract preparation procedure. Irrespective of whether or not Rim4 aggregates are disassembled prior to their degradation, the presence of amyloid-like Rim4 aggregates in cells correlated with Rim4-mediated translational repression. Clb3 protein levels accumulated as Rim4 aggregates were degraded (Figures 4A–4C). This finding raises the interesting possibility that it is the amyloid-like form of Rim4 that mediates translational repression.

To further explore whether translational repression of *CLB3* correlates with Rim4 aggregation, we utilized a construct where

Rim4 aggregates formed as soon as the protein was produced during premeiotic S phase and persisted throughout meiosis I. During meiosis II, SDS-resistant aggregates were no longer detected. We note that while SDS-resistant Rim4 aggregates were absent during meiosis II, we did detect Rim4 protein in extracts prepared by boiling cells in SDS-PAGE buffer (compare Figures 4A and 4B). We can envision two not mutually exclusive explanations for this observation. First, incomplete synchrony of cells progressing through meiosis could be responsible. The Rim4 protein detected in meiosis II time points by SDS-PAGE

CLB3 expression can be induced at will from the *GAL1-10* promoter (driven by a Gal4-ER fusion) but that preserves translational control of *CLB3* (henceforth *pGAL-5'UTR_{CLB3}-CLB3*). If Rim4 amyloid-like aggregation is important for its function, *CLB3* translational repression should correlate with aggregate formation. To test this prediction, we assessed when during early stages of sporulation *CLB3* becomes translationally repressed. We induced expression of *CLB3* from the *pGAL-5'UTR_{CLB3}-CLB3* fusion at various times after the induction of sporulation in cells lacking *NDT80* that progressed through sporulation in a synchronous manner (for a detailed description of the experimental setup see Figure S4A). *CLB3* translation occurred efficiently within the first 2 hr after induction of sporulation. By 3 hr, however, when cells had entered pre-meiotic S phase, *CLB3* translation was repressed as judged by the lack of Clb3 protein despite ample induction of *CLB3* expression (Figures 4D and S4B). Translational repression persisted until the end of the experiment when cells had arrested in G2 due to the lack of *NDT80*. Loss of *CLB3* translation correlated remarkably well with the formation of SDS-resistant Rim4 aggregates (Figures 4E and 4F). Taken together, these data indicate that cells generate RNA-binding amyloid-like aggregates of Rim4 to mediate translational control of gene expression in vivo.

The LC Sequences of Rim4 Mediate Formation of Amyloid-like Aggregates and Translational Repression

To determine whether the formation of amyloid-like Rim4 aggregates was required for translational repression of *CLB3* mRNA, we first identified the region within Rim4 responsible for aggregation in vivo. Two recent reports have shown that RNA stimulates formation of amyloid-like fibers of the RNA-binding protein FUS by increasing local concentration and thus nucleating assembly (Kato et al., 2012; Schwartz et al., 2013). Does Rim4's RNA-binding property mediate amyloid-like aggregation? To address this question, we analyzed SDS-resistant aggregate formation in a *rim4-F139L* mutant. The *rim4-F139L* mutant encodes a protein in which a critical phenylalanine residue in the first RRM domain is mutated to leucine (Soushko and Mitchell, 2000). The RRM mutant formed amyloid-like aggregates equally well as wild-type Rim4 in vitro and in vivo. Recombinant Rim4-F139L purified and concentrated to 27 mg/ml forms fibers in vitro (Figure 5A) and SDS-resistant aggregates in vivo (Figure 5B). We conclude that RNA-binding is not a prerequisite for aggregation in vitro and in vivo.

The LC sequences of Rim4 are a prime candidate for mediating the formation of Rim4 amyloid-like aggregates. Indeed, deleting the C-terminal 271 amino acids that comprise the LC sequences (Figure 5C) abrogated the formation of SDS-resistant aggregates in vivo (Figure 4A). As an alternative method to assess the importance of the C-terminal region of Rim4 in aggregate formation, we employed differential centrifugation. The unusual detergent-insolubility of amyloid-like aggregates allows for the sedimentation of these particles by ultracentrifugation from native lysates in the presence of detergent (Sondheimer and Lindquist, 2000). We assayed the sedimentation properties of *RIM4-3V5* as well as two *rim4* mutants lacking either most (*rim4Δ271C-3V5*) or a portion of the LC sequences

(*rim4Δ138C-3V5*). We observed full-length Rim4 protein predominantly in the pellet fraction while the aggregate-defective mutants remained in the supernatant (Figure 5D). We conclude that the C-terminal LC sequences of Rim4 are required for aggregation of the protein in vivo.

If amyloid-like aggregates are required for translational control, then abrogation of Rim4 aggregation should result in loss of *CLB3* translational repression. To test this prediction, we examined the consequences of deleting increasingly larger portions of the C-terminal LC sequences of Rim4 on aggregate formation and *CLB3* translation. We arrested cells in meiotic G2 by deleting *NDT80* and then induced expression of *CLB3* from the aforementioned *pGAL-5'UTR_{CLB3}-CLB3* fusion. In cells expressing full-length Rim4, *CLB3* translation was inhibited (Figures 5C and 5E). However, deleting as few as the C-terminal 66 amino acids led to *CLB3* translation in G2-arrested cells (Figures 5C and 5E). Larger deletions further increased translation efficiency.

Expressing Clb3 in G2-arrested cells had consequences. Upon *CLB3* induction, *rim4* aggregation-defective cells inappropriately formed meiosis I spindles in the G2 arrest, causing the appearance of cells with characteristics of both the *ndt80Δ* G2 arrest and meiosis I. Cells harbored separated spindle pole bodies (a meiosis I characteristic), and the synaptonemal complex protein Zip1 accumulated in the nucleus (a *ndt80Δ* arrest characteristic; Figure S5). Importantly, Rim4 truncations that failed to inhibit *CLB3* translation also failed to form SDS-resistant aggregates (Figure 5F). These findings show that aggregation and inhibition of translation are mediated by Rim4's LC domain.

Because the larger C-terminal deletions led to an ~50% decrease in Rim4 protein levels (Figure 5E), we also examined the effects of reducing the levels of full-length Rim4 by a similar degree. A *RIM4/rim4Δ* heterozygous strain produces approximately equivalent amounts of Rim4 as a strain carrying two copies of a *RIM4* allele producing a truncation lacking the C-terminal 138 aa (*rim4Δ138C-3V5*) but translational repression was largely unaffected compared to *rim4Δ138C-3V5* cells (Figures 5E and S5). We conclude that the loss of translational repression observed in *rim4* aggregation-defective mutants is not due to decreased protein levels. Rather, we propose that it is caused by an inability to form high molecular weight aggregates.

Rim4 Amyloid-like Aggregates Sequester *CLB3* mRNA

Aggregation-defective *rim4* mutants could abrogate translational control because they fail to bind RNA, repress translation, or both. To determine whether amyloid-like aggregation of Rim4 is important for RNA-binding, we performed extract-binding studies. We transcribed the *CLB3* 5'UTR along with two controls in vitro: the *CLB3* 5'UTR with a 25 nucleotide (nt) deletion and the 248 nt *CLB1* 5'UTR. These controls bind Rim4 with decreased affinity compared to the *CLB3* 5'UTR (Berchowitz et al., 2013). We coupled these RNAs via a biotin group to streptavidin beads and assayed the ability of the RNA to pull down Rim4 or an aggregation-defective mutant (Rim4Δ138C) from extracts prepared from G2-arrested cells, when Rim4 inhibits translation. We found that the RNA-binding activity of

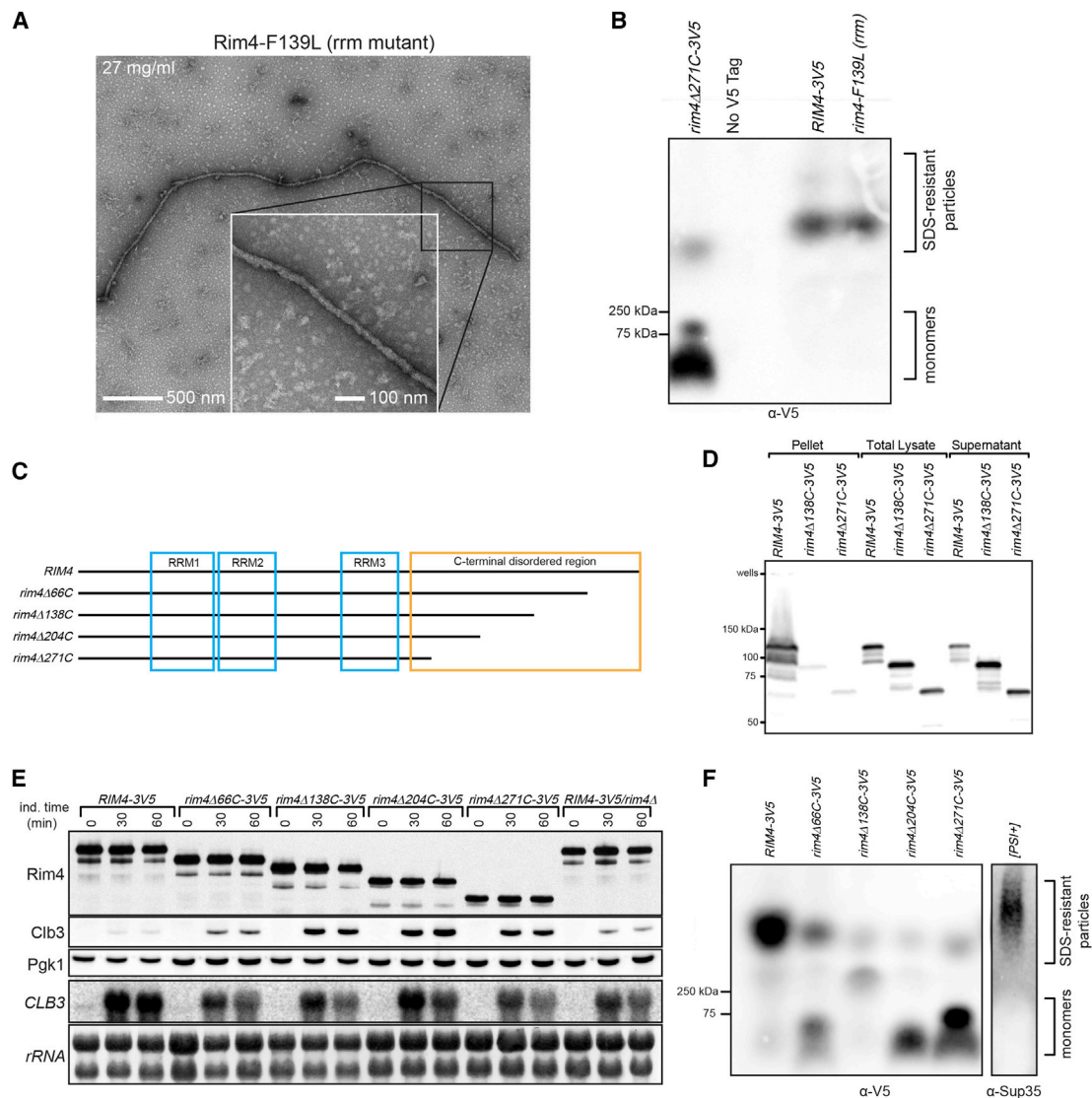


Figure 5. The LC Sequences, but Not RNA-Binding, Are Required for Rim4 Aggregation

(A) TEM of fibers formed by Rim4-F139L (RRM mutant) concentrated to 27 mg/ml and incubated at 4°C overnight.

(B) RNA binding is not required for amyloid-like aggregation. Rim4 was analyzed by SDD-AGE 6 hr after transfer into sporulation-inducing conditions of cells harboring *rim4Δ271C-3V5* (A35408), a no-tag control (A15055), *RIM4-3V5* (A35430), and *rim4-F139L-3V5* (RRM mutant A35324).

(C) Diagram of Rim4 C-terminal truncations.

(D) Rim4 can be pelleted from detergent-containing undenatured lysate. Strains harboring *RIM4-3V5* (A35430), *rim4Δ138C-3V5* (A35402), and *rim4Δ271C-3V5* (A35408) were arrested in G2 as above. Native lysates were prepared and spun for 60 min at 250,000 × g. Rim4 levels were determined in total lysate, pellet, and supernatant.

(E and F) *rim4* mutants that fail to form amyloid-like aggregates are defective for translational control. Strains harboring *RIM4-3V5* (A35430), heterozygous *RIM4-3V5/rim4Δ* (A35841), and constructs shown in (C) (A35399, A35402, A35405, A35408) as well as *pGAL-5'UTR_{CLB3}-CLB3*, *ndt80Δ*, *ZIP1-GFP*, *SPC42-mCherry*, and *GAL4.ER* were arrested in G2 (6 hr) and *pGAL-5'UTR_{CLB3}-CLB3* was induced by β-estradiol addition. Rim4-3V5, Clb3-3HA, and Pgk1 (loading control) protein and *CLB3* mRNA and *rRNA* (loading control) were determined (E). SDS-resistant Rim4 aggregates (at 6 hr) were analyzed in (F). The Sup35 aggregate-forming strain ([PSI⁺]; A25208) was used as a positive control.

See also Figure S5 and Table S2.

the aggregation-defective Rim4 was reduced by 50% (Figures 6A and 6B). These results indicate that the ability of Rim4 to aggregate is important for its ability to bind RNA. However, the 2-fold reduction in RNA-binding exhibited by Rim4Δ138C is unlikely to explain the >20-fold increase in translational efficiency of

the mutant. We propose that while amyloid-like aggregation is required for efficient RNA-binding, Rim4 aggregates participate in translational control beyond their role in RNA-binding.

Having determined that Rim4 aggregation plays only a minor role in RNA binding, we next tested the possibility that the

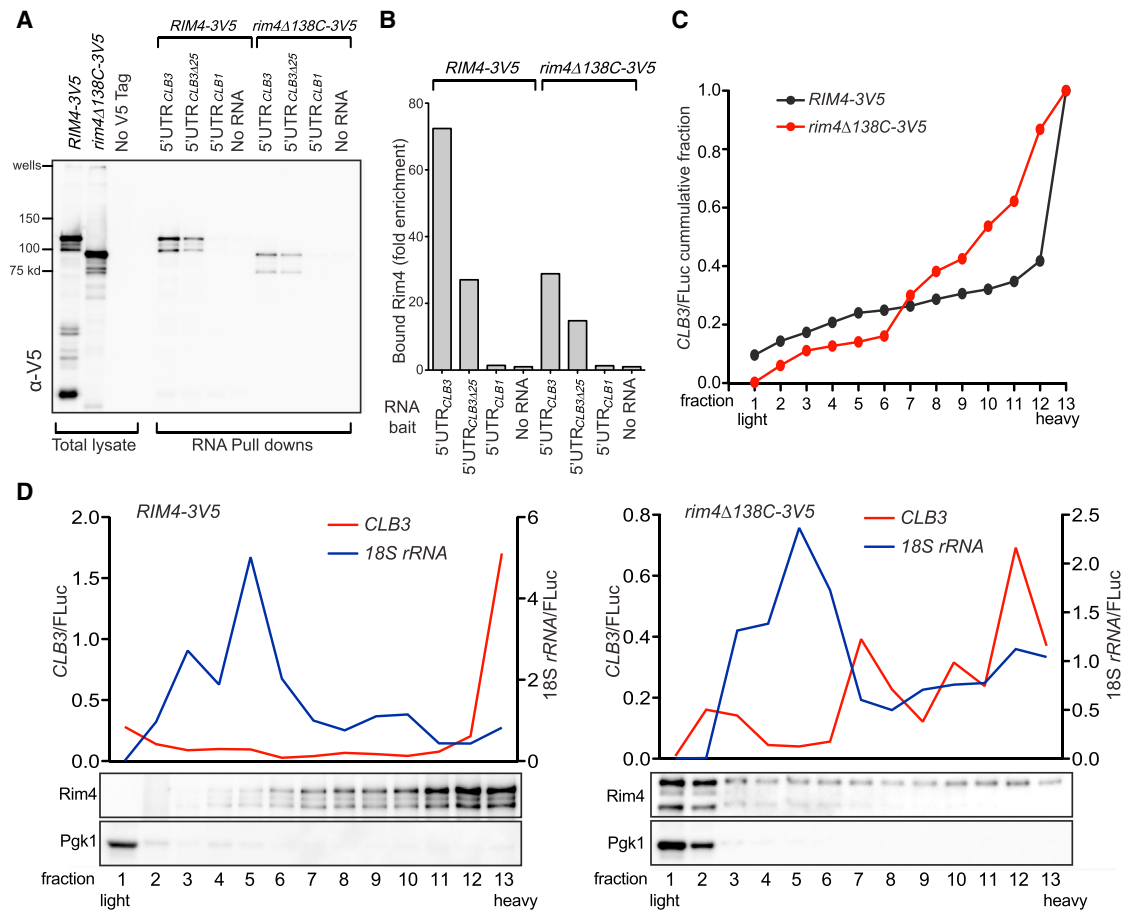


Figure 6. Effects of the LC Sequences on Rim4 RNA Binding and CLB3 mRNA Mobility in Sucrose Gradients

(A and B) An aggregation-defective *rim4* mutant has decreased RNA-binding activity. Extracts from *RIM4-3V5* (A35430) and *rim4Δ138C-3V5* (A35402) cells arrested in G2 (6 hr) were incubated with in vitro transcribed 3' biotinylated RNAs conjugated to streptavidin magnetic beads. Beads were recovered and boiled in SDS loading buffer to release bound factors. Shown is the amount of Rim4 recovered using the indicated RNA baits (A). Quantifications (B) were input-normalized and are shown as fold enrichment over the no-RNA control.

(C and D) Rim4 and its bound mRNA targets exist in heavy mRNPs. Strains harboring *RIM4-3V5* (A35430) and *rim4Δ138C-3V5* (A35402) were arrested in G2 as above except *CLB3* was expressed by the addition of β -estradiol for 60 min prior to collection. Samples were fixed in 1% formaldehyde on ice for 15 min (stopped by the addition of 0.1 M glycine) to preserve the RNPs (Valásek et al., 2007). Native lysates were fractionated on 10%–50% sucrose gradients. Sedimentation of Rim4, Pgk1, as well as *CLB3* mRNA and 18S rRNA was determined by immunoblot and qPCR, respectively. RNA quantification was normalized to luciferase RNA that was spiked into each fraction in equal amounts to normalize for RNA extraction efficiency.

See also Table S2.

Rim4 amyloid-like aggregates mediate repression of *CLB3* mRNA translation. If Rim4 aggregates repress translation, *CLB3* mRNA should cofractionate with Rim4 aggregates but not Rim4 monomers or oligomers. Furthermore, in *rim4* mutants that fail to form amyloid-like aggregates, *CLB3* mRNA ought to fractionate with translating ribosomes. To test these predictions, we fractionated lysates from *RIM4-3V5* and *rim4Δ138C-3V5* cells arrested in meiotic G2 (when *CLB3* is translationally repressed) in 10%–50% sucrose gradients. This analysis showed that both Rim4 and *CLB3* mRNA were present in the heaviest fractions of the gradient (Figures 6C and 6D). In contrast, Rim4Δ138C was predominantly present in the light fractions of the gradient and did not co-fractionate with *CLB3* mRNA. Instead, *CLB3* mRNA was found enriched in heavy frac-

tions of the gradient likely due to its association with translating ribosomes. Our analyses of wild-type and Rim4 mutant proteins in SDD-AGE as well as sedimentation and sucrose gradients indicate that it is the aggregated form of Rim4 that mediates translational repression.

Gametogenesis Signals Regulate Rim4 Amyloid Formation

Our results show that Rim4 forms amyloid-like aggregates during meiosis I to inhibit *CLB3* translation. Is this process regulated, and if so, which aspect of the regulatory network that induces sporulation controls the aggregation of Rim4? Two signals trigger gametogenesis in budding yeast: mating type and starvation (reviewed in van Werven and Amon, 2011). Only cells

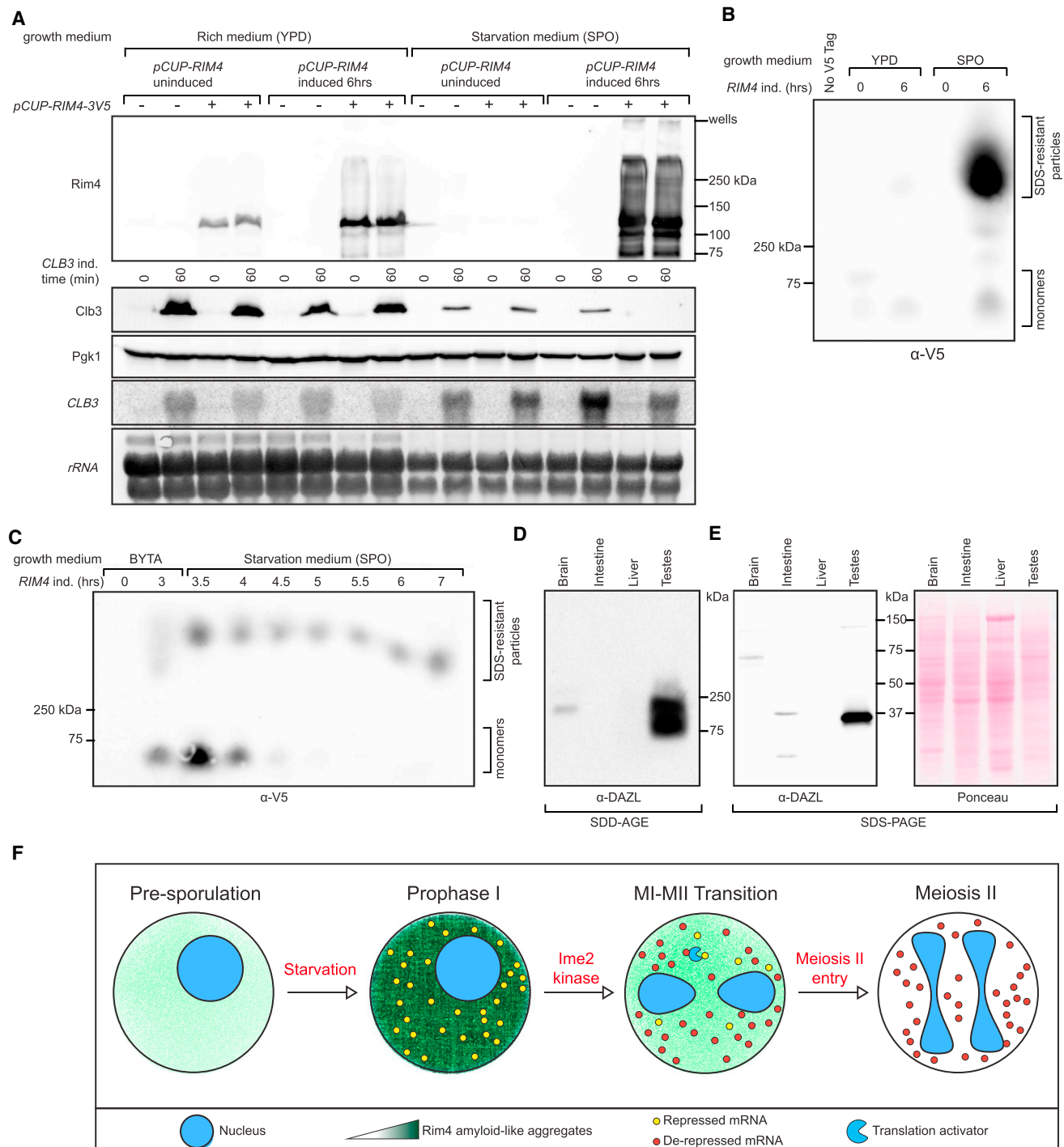


Figure 7. Formation of Rim4 Amyloid-like Aggregates Is Regulated by Starvation

(A and B) Starvation triggers Rim4 aggregation and translational repression (for detailed workflow see Figure S6A). Haploid *GAL4.ER*, *pGAL-5'UTR_{CLB3}-CLB3* strains with or without *pCUP1-RIM4-3V5* (A34980, A28492) were diluted to $OD_{600} = 1.8$ in either YEPD or SPO medium. *pCUP1-RIM4* was induced by the addition of $25 \mu\text{M}$ CuSO_4 . *pGAL-5'UTR_{CLB3}-CLB3* was induced by $2 \mu\text{M}$ β -estradiol prior to (0 hr) and after (6 hr) CuSO_4 induction for 60 min. Rim4-3V5, Clb3-3HA, and Pgk1 (loading control) protein and *CLB3* mRNA and *rRNA* (loading control) levels were determined at the indicated times in (A). SDS-resistant Rim4 aggregates (prior to and after CuSO_4 induction) were analyzed by SDD-AGE in (B).

(C) Monomeric Rim4 is converted into amyloid-like aggregates upon starvation (for detailed workflow see Figure S6B). A haploid *pGAL-RIM4-3V5* strain was grown in BYTA medium in the presence of $1 \mu\text{M}$ β -estradiol for 3 hr to induce Rim4 expression. Cells were then shifted to SPO medium. Rim4 SDS-resistant aggregates were assayed. Rim4 and Pgk1 protein levels from TCA extracts were analyzed in Figure S6C.

(legend continued on next page)

of the MATa/ α mating type can undergo sporulation and they do so only when they are starved, that is when they are deprived of nitrogen and a fermentable carbon source.

To determine whether formation of Rim4 amyloid-like aggregates is regulated by these gametogenesis-inducing signals, we asked whether ectopically expressed Rim4 aggregates in haploid cells grown in nutrient-rich medium supplied with ample nitrogen and glucose (YEPD medium). *RIM4* expressed under the control of the copper-inducible *CUP1* promoter readily produced Rim4 protein in YEPD medium but the protein did not form amyloid-like aggregates as judged by the lack of SDS-resistant Rim4 particles (Figures 7A, 7B, and S6A). Accordingly, *CLB3* was translated effectively (Figure 7A). In contrast, starvation caused Rim4 to form SDS-resistant aggregates. We starved haploid cells for 4 hr in sporulation medium and then induced Rim4 expression. Rim4 readily aggregated under these growth conditions and inhibited translation of *CLB3* (Figures 7A and 7B). These results indicate that the gametogenesis-inducing starvation signal causes the formation of Rim4 amyloid-like aggregates. Our data further demonstrate that monomeric Rim4 cannot repress *CLB3* translation.

To directly test that starvation induces Rim4 aggregation, we expressed Rim4 in haploid cells grown in nutrient rich conditions (BYTA) when the protein is largely in its monomeric form. We then repressed *RIM4* expression and simultaneously starved cells by transferring them into sporulation medium (Figures 7C and S6B). Within 30 min of inducing starvation, monomeric Rim4 converted into amyloid-like aggregates (Figures 7C and S6C). We conclude that aggregation of Rim4 is a regulated process. Monomeric Rim4 assembly into amyloid-like aggregates is induced by starvation signals to convert the protein into an amyloid-like translational repressor.

DAZL Forms SDS-Resistant Aggregates In Vivo

Although budding yeast harbors many proteins with RRM domains and a prion domain, only Rim4 is sporulation-specific. In humans, there are only four such gametogenesis-specific proteins. They are *DAZ1–DAZ4* (King et al., 2012). These four proteins exist as a cluster on the Y-chromosome- a region of the Y that is often found deleted in infertile men with a condition called azoospermia (Reijo et al., 1995). While *DAZ* exists only in humans and old world monkeys, one of its autosomal homologs *DAZL* (*DAZ*-like) is found from fish to humans (Haag, 2001). Like Rim4, *DAZL* functions in the regulation of translation (Collier et al., 2005) and is essential for gametogenesis in many organisms, including mice (Ruggiu et al., 1997). We found that mouse *DAZL* forms SDS-resistant aggregates in cells undergoing gametogenesis (testes), but not in brain, intestine, and liver (Figures 7D and 7E). We propose that *DAZL* (and possibly *DAZ*) functions to regulate translation of gametogenesis-specific RNAs in a manner similar to Rim4 and suggest that amyloid-like aggregation of RNA-binding proteins is an evolutionarily conserved feature of gametogenesis.

DISCUSSION

Here, we show that amyloid-like aggregates, generally considered to be inactive and often toxic, play a central role in creating the cell division essential for gamete formation and hence sexual reproduction. Rim4 forms amyloid-like aggregates to repress the translation of *CLB3* during meiosis I that is critical to establish the meiosis I chromosome segregation pattern. Importantly, formation of these structures is regulated. Starvation, one of the signals that triggers gamete development, induces the conversion of monomeric Rim4 into aggregated Rim4. Thus, starvation not only initiates gametogenesis but also sets up the mechanism that brings about the meiotic chromosome segregation pattern much later during gamete formation.

Rim4 Aggregation Is a Regulated Process

Several lines of evidence demonstrate that Rim4 forms large amyloid-like aggregates in vivo. The protein is immobile and found in granular structures within cells, forms SDS-resistant aggregates when isolated from cell extracts, can be sedimented and fractionated from undenatured detergent-containing lysates, and can be quantitatively precipitated from cell extracts using b-isox. The protein's in vitro properties: Thioflavin T binding, fiber formation, CD spectrum, and reactivity with an anti-amyloid antibody further support the idea that Rim4 has amyloid-like properties.

Our data lead to the conclusion that aggregation of Rim4 is a regulated process. When we express Rim4 in cells grown in nutrient-rich conditions the protein is monomeric, but upon transfer into sporulation-inducing conditions the protein quickly converts into the aggregated form. During normal progression through sporulation, we do not observe this conversion. This is not surprising because *RIM4* is only induced during premeiotic DNA replication, after cells have been transferred into starvation-inducing medium. Thus, Rim4 is immediately assembled into amyloid-like aggregates upon its expression. We do not yet know how starvation induces Rim4 assembly. Downregulation of the RAS and TOR pathways is likely to play a central role in the process. Once we understand how amyloid assembly is controlled in vivo, we can begin to think about strategies to reverse the process. In time, these could very well result in the identification of anti-amyloid treatments.

Once Rim4 aggregates into amyloid-like structures and binds to target RNAs, it inhibits translation until the completion of meiosis I (Figure 7F). At the onset of meiosis II, Rim4 aggregates are degraded by an unknown mechanism allowing translation to commence. Our studies suggest that the C-terminal LC sequences are critical for this degradation. This finding raises the interesting possibility that Rim4 must be in the amyloid-like form to be cleared from cells. This would imply that autophagy degrades Rim4 structures at the end of meiosis I. However, we have not detected Rim4-EGFP in the vacuole at the time of its degradation at the onset of meiosis II, arguing against autophagy-mediated clearance of Rim4. Rim4 could be eliminated

(D and E) *DAZL* forms SDS-resistant aggregates in mouse gametogenesis. Brain, intestine, liver, and testes were harvested from a 2-month-old C57BL/6J male mouse. Samples were Dounce homogenized in SDD-AGE buffer and analyzed by SDD-AGE (D) and SDS-PAGE (E) *DAZL* was detected by immunoblot analysis. (F) Model for Rim4-mediated translational control. See text for details. See also Table S2.

by proteasomal degradation, but such a mechanism would likely require prior aggregate disassembly. Our previous studies have implicated Ime2-mediated phosphorylation in triggering Rim4 degradation (Berchowitz et al., 2013). Perhaps Ime2 phosphorylation triggers Rim4 disassembly and degradation by the proteasome. Amyloids and amyloid-like aggregates are usually stable and long-lived. Understanding how they are cleared from cells at the appropriate developmental transition leaving Rim4 target RNAs intact ready to be translated will be critically important.

Mechanisms of Translational Repression by Rim4

Why must Rim4 assemble into amyloid-like particles to repress translation? An aggregate-based apparatus rather than a mechanism involving monomer binding could be used to localize the translationally inhibited mRNA to certain sites within the cell where the protein products are eventually needed. This is unlikely to be the case for Rim4. The protein does not show any localization preference within the cytoplasm (Berchowitz et al., 2013). The amyloid-like state could allow for heterotypic interactions with other proteins with amyloid-like properties that are involved in translational repression. Amyloids have the ability to influence cross-seeding of other amyloidogenic proteins (Derkatch and Lieberman, 2007). A computational survey in budding yeast identified 179 proteins harboring prion domains many of which exhibit canonical properties of amyloid-forming proteins (Alberti et al., 2009). Many of these proteins are involved in RNA processing and translational control. Cross-seeding between proteins with amyloid-like properties could offer a mechanism for the assembly of translationally repressive cytoplasmic structures similar to P-bodies (Decker et al., 2007).

Translational Control Mediated by Amyloid-like Aggregates Could Be a Universal Property of Gametogenesis

A significant proportion of RRM-containing proteins also contain a prion domain and/or LC sequences. This raises the interesting possibility that regulation of RNA functions by amyloid-like aggregates is a frequently used mechanism of controlling gene expression (Kato et al., 2012; King et al., 2012). In budding yeast, ~1% of genes encode proteins bearing RRM domains (52 genes) and ~3% of genes encode proteins containing a prion domain (179 genes) (Alberti et al., 2009). Thirteen genes encode proteins with both an RRM and prion domain. This is ~10-fold higher than is expected to occur at random. A similar situation exists in humans. The coexistence of RRM and prion domains within the same protein is 13-fold higher than what would be expected to occur by chance (King et al., 2012). This association has led to the proposal that multimerization via prion domains or LC sequences is necessary for the function of many RNA-binding proteins. Indeed, several RNA-binding proteins containing a prion domain have well-characterized roles in translational control such as Ccr4, Pat1, TIF4632 (eIF4G), and Whi3. Understanding under which circumstances an amyloid-like structure can promote or repress translation will provide critical insights into the pathogenesis of amyloids.

Remarkably, both yeast and mouse contain an RNA-binding protein, Rim4 and DAZL, respectively, that functions in translation and forms SDS-resistant aggregates during gametogenesis.

This finding raises the possibility that DAZL forms amyloid-like aggregates to regulate translation. Additionally, it leads to the speculation that gametogenic processes as diverse as spore formation in yeast and sperm production in mammals are fundamentally related and rely on an ancient process that involves the formation of large amyloid-like aggregates.

EXPERIMENTAL PROCEDURES

Sporulation Conditions

Strains were grown to saturation in YPD, diluted in BYTA (1% yeast extract, 2% tryptone, 1% potassium acetate, 50 mM potassium phthalate) to OD₆₀₀ = 0.25 and grown overnight. Cells were resuspended in sporulation medium (0.3% potassium acetate [pH 7.0], 0.02% raffinose) to OD₆₀₀ = 1.8 and sporulated with vigorous shaking at 30°C. *pGAL-NDT80*, *GAL4.ER* strains were released from the arrest by the addition of 1 μM β-estradiol at 6 hr. In *ndt80Δ* strains, expression of *pGAL-5'UTR_{CLB3}CLB3-3HA* was induced with 1 μM β-estradiol at the indicated times.

Transmission Electron Microscopy

For negative staining, 10 μl of protein solution was applied to a TEM grid (Electron Microscopy Sciences), washed with distilled water, and stained with several drops of 1% or 2% (w/v) aqueous uranyl acetate. Excess stain was removed and the grids air-dried. TEM images were obtained at 80 kV on an FEI Technai Spirit TEM. For cryo-TEM, 10 μl of protein solution was applied to a TEM grid, washed with distilled water, and plunged into cooled liquid ethane. TEM images were obtained at 120 kV on a JEOL 2100 FEG TEM.

Semi-denaturing Detergent Agarose Gel Electrophoresis

Samples (2 ml) of sporulating culture were harvested by centrifugation and snap frozen in liquid nitrogen. Cell pellets were resuspended in lysis buffer (100 mM Tris/HCl pH 8.0, 20 mM NaCl, 2 mM MgCl₂, 50 mM β-mercaptoethanol, 1% Triton-X, 2× Halt protease inhibitors [Thermo]) and broken using a Fast Prep (MP bio) and zirconia beads (BioSpec). The lysates were clarified twice by centrifugation at 2,500 relative centrifugal force (rcf) for 5 min (4°C). Loading buffer (final concentration 0.5× TAE, 5% glycerol, 2% SDS, bromophenol blue) was added to lysates that were then incubated for 10 min at room temperature. Following incubation, samples were separated on a 1.7% agarose gel for 16 hr at 29 V at 4°C using 1× TAE supplemented with 0.1% SDS as the running buffer. The gel was blotted to nitrocellulose by downward capillary transfer in 1× TBS and processed by the immunoblot procedure outlined in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, two tables, and six movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.060>.

ACKNOWLEDGMENTS

We thank Steven McKnight and Doug Frantz for providing the b-isox reagent and Susan Lindquist for providing the α-Sup35 antibody. We thank Kristin Knouse for help with DAZL analysis, Arzu Sandikçi for development and troubleshooting microfluidics and help with live imaging analysis, and Matthew Webber for help with CD. We thank Amanda Del Rosario for mass spectrometry and Dong Soo Yun for TEM at the Koch Cancer Institute Core Facilities. We thank Wendy Salmon and Nikki Watson at the Whitehead Institute for help with FRAP and TEM. We thank members of the Amon lab for critical reading of the manuscript. All animal procedures were approved by the Massachusetts Institute of Technology Committee on Animal Care. L.E.B. is supported by American Cancer Society and Charles King Trust post-doctoral fellowships. G.L.K. is supported by an NIH grant to T.U.S. T.M.C. is supported by an American Cancer Society fellowship. This research was supported by NIH grants GM62207 to A.A., GM77537 to T.U.S., and GM094303 to W.V.G. A.A. is also an investigator of the Howard Hughes Medical Institute.

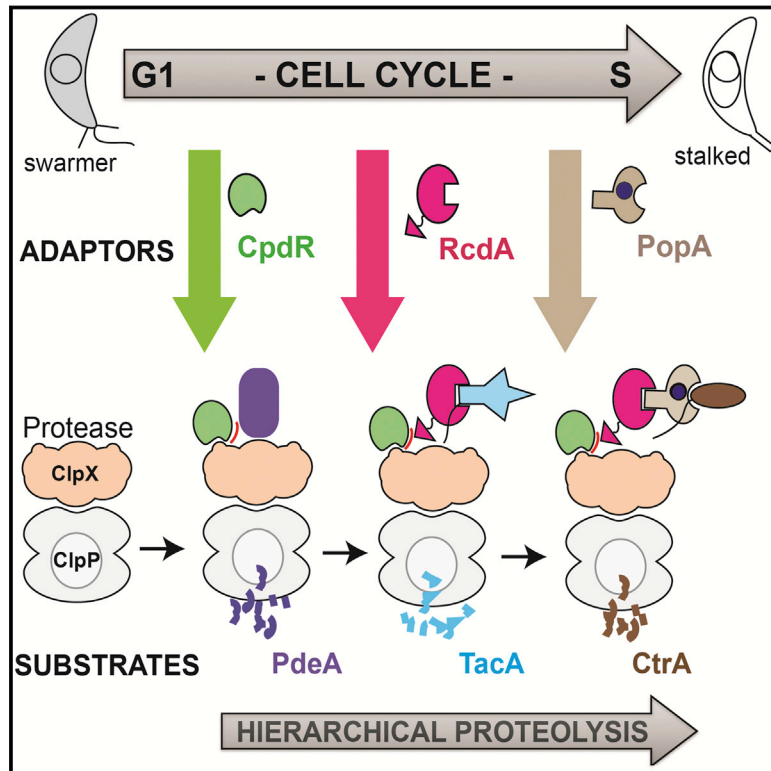
Received: March 1, 2015
Revised: June 28, 2015
Accepted: August 17, 2015
Published: September 24, 2015

REFERENCES

- Akiyama, H., Barger, S., Barnum, S., Bradt, B., Bauer, J., Cole, G.M., Cooper, N.R., Eikelenboom, P., Emmerling, M., Fiebich, B.L., et al. (2000). Inflammation and Alzheimer's disease. *Neurobiol. Aging* 21, 383–421.
- Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158.
- Benjamin, K.R., Zhang, C., Shokat, K.M., and Herskowitz, I. (2003). Control of landmark events in meiosis by the CDK Cdc28 and the meiosis-specific kinase Ime2. *Genes Dev.* 17, 1524–1539.
- Berchowitz, L.E., Gajadhar, A.S., van Werven, F.J., De Rosa, A.A., Samoylova, M.L., Brar, G.A., Xu, Y., Xiao, C., Fletcher, B., Weissman, J.S., et al. (2013). A developmentally regulated translational control pathway establishes the meiotic chromosome segregation pattern. *Genes Dev.* 27, 2147–2163.
- Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416, 507–511.
- Carlile, T.M., and Amon, A. (2008). Meiosis I is established through division-specific translational control of a cyclin. *Cell* 133, 280–291.
- Collier, B., Gorgoni, B., Loveridge, C., Cooke, H.J., and Gray, N.K. (2005). The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. *EMBO J.* 24, 2656–2666.
- Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J. Cell Biol.* 179, 437–449.
- DePace, A.H., Santoso, A., Hillner, P., and Weissman, J.S. (1998). A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93, 1241–1252.
- Derkatch, I.L., and Liebman, S.W. (2007). Prion-prion interactions. *Prion* 1, 161–169.
- Fowler, D.M., Koulov, A.V., Alory-Jost, C., Marks, M.S., Balch, W.E., and Kelly, J.W. (2006). Functional amyloid formation within mammalian tissue. *PLoS Biol.* 4, e6.
- Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell.* 15, 5383–5398.
- Haag, E.S. (2001). Rolling back to BOULE. *Proc. Natl. Acad. Sci. USA* 98, 6983–6985.
- Halfmann, R., Jarosz, D.F., Jones, S.K., Chang, A., Lancaster, A.K., and Lindquist, S. (2012). Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature* 482, 363–368.
- Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767.
- Kayatekin, C., Matlack, K.E.S., Hesse, W.R., Guan, Y., Chakrabortee, S., Russ, J., Wanker, E.E., Shah, J.V., and Lindquist, S. (2014). Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. *Proc. Natl. Acad. Sci. USA* 111, 12085–12090.
- Kedersha, N.L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 α to the assembly of mammalian stress granules. *J. Cell Biol.* 147, 1431–1442.
- King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* 1462, 61–80.
- Knowles, T.P.J., Vendruscolo, M., and Dobson, C.M. (2014). The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* 15, 384–396.
- Kronja, I., and Orr-Weaver, T.L. (2011). Translational regulation of the cell cycle: when, where, how and why? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 3638–3652.
- Li, J., McQuade, T., Siemer, A.B., Napetschnig, J., Moriwaki, K., Hsiao, Y.S., Damko, E., Moquin, D., Walz, T., McDermott, A., et al. (2012). The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis. *Cell* 150, 339–350.
- Loo, D.T., Copani, A., Pike, C.J., Whittemore, E.R., Walencewicz, A.J., and Cotman, C.W. (1993). Apoptosis is induced by beta-amyloid in cultured central nervous system neurons. *Proc. Natl. Acad. Sci. USA* 90, 7951–7955.
- Maji, S.K., Perrin, M.H., Sawaya, M.R., Jessberger, S., Vadodaria, K., Rissman, R.A., Singru, P.S., Nilsson, K.P., Simon, R., Schubert, D., et al. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325, 328–332.
- Majumdar, A., Cesario, W.C., White-Grindley, E., Jiang, H., Ren, F., Khan, M.R., Li, L., Choi, E.M., Kannan, K., Guo, F., et al. (2012). Critical role of amyloid-like oligomers of *Drosophila* Orb2 in the persistence of memory. *Cell* 148, 515–529.
- Miller, M.P., Unal, E., Brar, G.A., and Amon, A. (2012). Meiosis I chromosome segregation is established through regulation of microtubule-kinetochore interactions. *eLife* 1, e00117.
- Nilsson, M.R. (2004). Techniques to study amyloid fibril formation in vitro. *Methods* 34, 151–160.
- Reijo, R., Lee, T.Y., Salo, P., Alagappan, R., Brown, L.G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., Hovatta, O., et al. (1995). Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat. Genet.* 10, 383–393.
- Dos Reis, S., Couлары-Salin, B., Forge, V., Lascu, I., Bégueret, J., and Saupe, S.J. (2002). The HET-s prion protein of the filamentous fungus *Podospora anserina* aggregates in vitro into amyloid-like fibrils. *J. Biol. Chem.* 277, 5703–5706.
- Ruggiu, M., Speed, R., Taggart, M., McKay, S.J., Kilanowski, F., Saunders, P., Dorin, J., and Cooke, H.J. (1997). The mouse Dazl gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* 389, 73–77.
- Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G.P., Davies, S.W., Lehrach, H., and Wanker, E.E. (1997). Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90, 549–558.
- Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA seeds higher-order assembly of FUS protein. *Cell Rep.* 5, 918–925.
- Sipe, J.D., and Cohen, A.S. (2000). Review: history of the amyloid fibril. *J. Struct. Biol.* 130, 88–98.
- Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. *Mol. Cell* 5, 163–172.
- Soushko, M., and Mitchell, A.P. (2000). An RNA-binding protein homologue that promotes sporulation-specific gene expression in *Saccharomyces cerevisiae*. *Yeast* 16, 631–639.
- Thomas, T., Thomas, G., McLendon, C., Sutton, T., and Mullan, M. (1996). β -Amyloid-mediated vasoactivity and vascular endothelial damage. *Nature* 380, 168–171.
- Valášek, L., Szamecz, B., Hinnebusch, A.G., and Nielsen, K.H. (2007). In vivo stabilization of preinitiation complexes by formaldehyde cross-linking. *Methods Enzymol.* 429, 163–183.
- van Werven, F.J., and Amon, A. (2011). Regulation of entry into gametogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 3521–3531.
- Walsh, D.M., and Selkoe, D.J. (2007). A beta oligomers - a decade of discovery. *J. Neurochem.* 101, 1172–1184.
- Xu, L., Ajimura, M., Padmore, R., Klein, C., and Kleckner, N. (1995). NDT80, a meiosis-specific gene required for exit from pachytene in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 15, 6572–6581.

An Adaptor Hierarchy Regulates Proteolysis during a Bacterial Cell Cycle

Graphical Abstract



Authors

Kamal Kishore Joshi, Matthieu Bergé,
Sunish Kumar Radhakrishnan,
Patrick Henri Viollier, Peter Chien

Correspondence

pchien@biochem.umass.edu

In Brief

Protein degradation during the *Caulobacter* cell cycle relies on an adaptor hierarchy. Priming of the protease by one adaptor recruits additional adaptors, and substrate degradation occurs based on the degree of adaptor assembly.

Highlights

- Protease selectivity during cell cycle is defined by a hierarchy of adaptors
- Priming of the ClpXP protease by CpdR is needed for adaptor recruitment
- RcdA is an adaptor that binds multiple substrates and additional adaptors
- Which substrates are degraded depends on the degree of adaptor assembly



An Adaptor Hierarchy Regulates Proteolysis during a Bacterial Cell Cycle

Kamal Kishore Joshi,¹ Matthieu Bergé,² Sunish Kumar Radhakrishnan,^{2,3} Patrick Henri Viollier,² and Peter Chien^{1,*}

¹Department of Biochemistry and Molecular Biology, Molecular and Cellular Biology Graduate Program, University of Massachusetts Amherst, Amherst, MA 01003, USA

²Department of Microbiology and Molecular Medicine, Institute of Genetics & Genomics in Geneva (iGE3), University of Geneva Medical School, Geneva CH-1211, Switzerland

³Present address: School of Biology, Indian Institute of Science Education and Research, Thiruvananthapuram, Kerala 695016, India

*Correspondence: pchien@biochem.umass.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.030>

SUMMARY

Regulated protein degradation is essential. The timed destruction of crucial proteins by the ClpXP protease drives cell-cycle progression in the bacterium *Caulobacter crescentus*. Although ClpXP is active alone, additional factors are inexplicably required for cell-cycle-dependent proteolysis. Here, we show that these factors constitute an adaptor hierarchy wherein different substrates are destroyed based on the degree of adaptor assembly. The hierarchy builds upon priming of ClpXP by the adaptor CpdR, which promotes degradation of one class of substrates and also recruits the adaptor RcdA to degrade a second class of substrates. Adding the PopA adaptor promotes destruction of a third class of substrates and inhibits degradation of the second class. We dissect RcdA to generate bespoke adaptors, identifying critical substrate elements needed for RcdA recognition and uncovering additional cell-cycle-dependent ClpXP substrates. Our work reveals how hierarchical adaptors and primed proteases orchestrate regulated proteolysis during bacterial cell-cycle progression.

INTRODUCTION

Regulated proteolysis is crucial for all life. For example, the timely destruction of key regulators by energy-dependent proteases during the eukaryotic and bacterial cell cycle drives replication and growth (King et al., 1996; Konovalova et al., 2014). Because proteolysis is irreversible, cells face the substantial challenge of stringently distinguishing specific proteins that are to be rapidly destroyed from many others meant to remain stable. In bacteria, energy-dependent proteases, such as the essential ClpXP protease, use accessory factors called adaptors to modulate substrate specificity (Gottesman, 2003; Ades, 2004; Guo and Gross, 2014). Adaptor proteins often work by binding directly to substrates and targeting them to appropriate proteases. One of the first characterized adaptors is SspB, which enhances ClpXP degradation of incomplete

translation products tagged by the ssrA peptide by binding to the peptide tag and tethering the substrate to the protease, catalyzing immediate destruction (Dogan et al., 2003; Levchenko et al., 2003). In this case, the adaptor enforces rapid degradation by increasing effective substrate concentration through tethering (Kirstein et al., 2009; Battesti and Gottesman, 2013a). Adaptor proteins can themselves be regulated by anti-adaptors, proteins that block adaptor activity during different growth or environmental conditions (Battesti and Gottesman, 2013a).

The Gram-negative α -proteobacterium *Caulobacter crescentus* has a dimorphic life cycle wherein an obligate transition from a swarmer cell stage to a stalked cell stage is driven by degradation of key regulatory proteins (Poindexter, 1981; Curtis and Brun, 2010). The essential transcription factor CtrA inhibits replication initiation and is degraded by ClpXP during the swarmer to stalked cell transition (Quon et al., 1996; Domian et al., 1997; Jenal and Fuchs, 1998). The destruction of CtrA promotes assembly of replication initiation machinery at the origin in stalked cells and drives the progression of the cell cycle (reviewed in Marczynski and Shapiro, 2002; Thanbichler, 2010). Similarly, the developmental regulator TacA, which activates transcription of genes involved in stalk biogenesis and polar development (Biondi et al., 2006; Radhakrishnan et al., 2008), is degraded by ClpXP in a cell-cycle-dependent manner (Bhat et al., 2013).

Degradation of both TacA and CtrA depends on the response regulator CpdR (Iniesta et al., 2006; Bhat et al., 2013). Previous reconstitution experiments have shown that CpdR enhances degradation of some ClpXP substrates in a phosphorylation-dependent manner by acting as an adaptor to activate substrate recognition (Abel et al., 2011; Rood et al., 2012; Lau et al., 2015). However, addition of the CpdR adaptor alone does not improve degradation of CtrA in vitro, but addition of two additional factors, RcdA and PopA, promotes degradation of CtrA by ClpXP in a second messenger and phosphorylation-dependent manner (Smith et al., 2014). RcdA and PopA were originally identified as necessary for cell-cycle-dependent CtrA degradation in vivo (McGrath et al., 2006; Duerig et al., 2009) and are thought to obligately interact with each other (Duerig et al., 2009; Ozaki et al., 2014). Interestingly, orthologs of CpdR and RcdA are found in most α -proteobacteria, but PopA is restricted to *Caulobacter* and closely related species (Brilli et al., 2010;

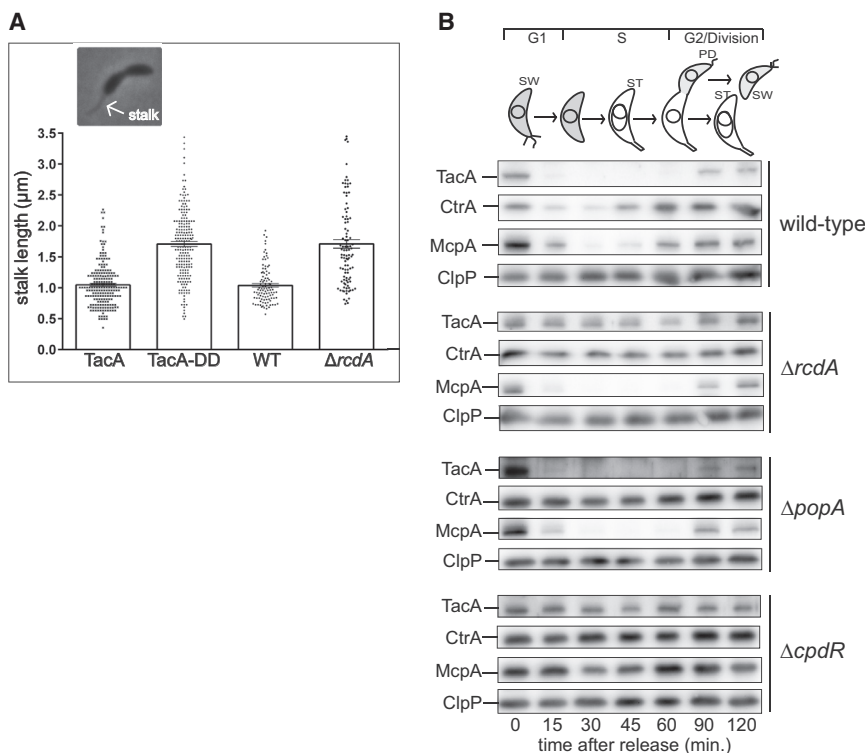


Figure 1. Cell-Cycle-Dependent Degradation of the TacA Substrate Requires RcdA and CpdR but Not PopA

(A) WT cells, $\Delta rcdA$ cells, and cells expressing either TacA or TacA-DD as a sole copy from the chromosomal locus were grown to exponential phase in PYE and imaged by phase-contrast microscopy. Lengths of visible stalks ($n = 104$ – 200) from a mixed population were measured using ImageJ. Each marker represents stalk lengths of an individual cell. Data represent mean \pm SEM. (B) WT cultures or cultures of cells lacking RcdA, CpdR, and PopA were synchronized, and equal numbers of swarmer populations were released into fresh PYE medium. Constant volume of samples were withdrawn at indicated time points and probed with anti-CtrA, anti-TacA, anti-McpA and anti-ClpP antibodies. ClpP, the levels of which remain constant throughout cell cycle, was used as a control. SW = swarmer cell, ST = stalk cell, PD = predivisional cell.

RESULTS

Degradation of the TacA Substrate Requires CpdR and RcdA but Not PopA

We previously identified TacA as a ClpXP substrate that requires CpdR for its cell-

cycle-dependent degradation (Bhat et al., 2013) (Figure 1B). During our characterization of this result, we found that expression of a nondegradable TacA allele (TacA-DD, Bhat et al., 2013) resulted in cells with longer stalks compared to wild-type (WT) cells (Figure 1A), consistent with the role of TacA in stalk biogenesis. Cells lacking RcdA were previously reported to have increased stalk length (McGrath et al., 2006), and quantitative measurements showed that these increased lengths were statistically indistinguishable from those in the TacA-DD strain (Figure 1A), suggesting that RcdA might be needed for proper TacA degradation. In accordance with this model, the cell-cycle-dependent degradation of TacA is lost in $\Delta rcdA$ strains (Figure 1B).

Prior work suggested that RcdA requires the PopA protein for its biological functions, such as the cell-cycle-dependent degradation of CtrA (Duerig et al., 2009) (Figure 1B). Indeed, the regulated degradation of CtrA has been recently reconstituted in vitro, and PopA was necessary to forming a complex with both RcdA and CtrA as part of this process (Smith et al., 2014). Furthermore, all ClpXP substrates that rely on RcdA for cell-cycle-dependent degradation also require PopA (Duerig et al., 2009; Radhakrishnan et al., 2010). Thus, we were surprised to find that although TacA degradation in vivo was dependent on RcdA, degradation was unaffected in a $\Delta popA$ strain (Figure 1B). These results show an unexpected role for RcdA separate from PopA in driving cell-cycle-dependent ClpXP proteolysis, a feature that we next explored biochemically.

Our original identification of TacA as a ClpXP substrate showed that it was slowly degraded by ClpXP in vitro and required CpdR for in vivo degradation (Bhat et al., 2013).

Because the CpdR adaptor delivers substrates such as PdeA to the ClpXP protease (Abel et al., 2011; Rood et al., 2012; Lau et al., 2015), we first tested whether CpdR could stimulate TacA degradation by ClpXP as well. However, TacA degradation was unchanged in the presence of CpdR alone (Figure 2A). We next asked whether addition of both RcdA and CpdR affected TacA degradation in vitro. Consistent with our in vivo observations (Figure 1B), TacA degradation by ClpXP was dramatically accelerated (>20-fold decrease in half-life) in the presence of both CpdR and RcdA (Figures 2A and S1B). This stimulation was specific as degradation of an unrelated substrate (GFP-ssrA) by ClpXP was unaffected by the addition of RcdA and CpdR (Figure S1C). RcdA and CpdR failed to deliver TacA with mutations on the extreme C-terminal residues to di-aspartate, suggesting that the native C terminus constitutes the protease recognition motif (Figure 2A) (Bhat et al., 2013). Taken together, these results confirm our in vivo findings that RcdA has a PopA-independent role in promoting degradation of certain CpdR-dependent ClpXP substrates.

The RcdA Adaptor Binds the TacA Substrate Directly to Promote Degradation

If CpdR and RcdA together make up a substrate-specific adaptor system for delivery of TacA to ClpXP, it seems reasonable that one of the two should interact directly with TacA. Based on size-exclusion chromatography, we found that the RcdA adaptor could directly bind to TacA (Figure 2B). TacA is an NtrC-family response regulator that contains an N-terminal receiver domain (RD), an ATP-binding AAA⁺ domain (AAA), and a C-terminal DNA-binding domain (DBD) (Figure 2C). To determine which region of TacA is needed for RcdA-dependent recognition and delivery, we generated a series of TacA truncations. Given that the degradation of the response regulator CtrA by CpdR/RcdA/PopA relies on its N-terminal receiver domain (Ryan et al., 2002), we expected a similar reliance for TacA. Surprisingly, the RD region was dispensable for adaptor-dependent proteolysis as fragments containing only the DBD bind to RcdA (Figure S1D) and are degraded similarly to full-length protein (Figure 2C). Further truncations and fusions revealed that a region of 12 residues (437–448) within TacA was necessary for RcdA binding (Figures 2C and 2D). In total, these results support a model wherein the C-terminal domain of TacA contains both RcdA and ClpXP recognition motifs that ensure robust degradation of TacA in a CpdR-dependent fashion (Figure 2E). We next sought to understand how the RcdA adaptor delivers the TacA substrate to the ClpXP protease and how this delivery depends on the adaptor CpdR.

The RcdA Adaptor Contains a Tethering Motif Needed for Substrate Delivery to the CpdR-Primed ClpXP Protease

Previous work had shown that the disordered C-terminal tail of the RcdA adaptor plays a critical role in the regulated degradation of the CtrA substrate in vivo (Taylor et al., 2009). To determine whether the C terminus of RcdA is also important for TacA degradation, we expressed a variant of RcdA lacking 19 residues from the C terminus (RcdAΔC) as the sole copy in vivo and monitored levels of TacA and CtrA during synchronous

growth. As expected, CtrA levels were stable throughout the cell cycle (Taylor et al., 2009) (Figure 3A). TacA levels also remain stable in this background, whereas the non-RcdA-dependent ClpXP substrate McpA was still degraded in a cell-cycle-dependent manner (Figure 3A). Consistent with these in vivo results, RcdAΔC did not stimulate degradation of TacA in vitro even in the presence of CpdR (Figure 3B). Interestingly, RcdAΔC is still capable of binding TacA (Figures 3C, S2A, and S2B), suggesting that the defect in TacA degradation is downstream of substrate binding. This architecture is reminiscent of the SspB adaptor, in which an N-terminal domain binds ssrA-tagged substrates and a ClpX-binding motif at the C terminus anchors the adaptor to the protease (Dougan et al., 2003; Levchenko et al., 2003). Therefore, we next sought to characterize how the C terminus of the RcdA adaptor contributes to substrate delivery.

The C terminus of RcdA harbors a number of hydrophobic residues with spacing consistent with an amphipathic helix (Figures 3D and S2C). Mutation of these residues inhibited RcdA-dependent delivery of the TacA substrate, consistent with a model in which the hydrophobic face of the putative amphipathic helix is needed for delivery. This result suggested that the RcdA tail acts as a binding element for CpdR, ClpX, or both as part of the mechanism by which the RcdA adaptor delivers substrates to ClpXP. Recently, we showed that CpdR acts as a priming adaptor, binding to the N-terminal domain of ClpX to enhance recruitment of substrates such as PdeA and McpA (Lau et al., 2015). We considered that the C-terminal tail of the RcdA adaptor might also be selectively binding to a CpdR-primed ClpXP protease, promoting substrate delivery.

We first tested this model by fusing the C-terminal 19 residues of RcdA to SspB lacking its ClpX-binding motif to generate a SspB~RcdA chimera (Figure 3E). Our rationale was that this chimeric construct would bind ssrA-tagged substrates similar to SspB but with the protease-binding specificity of RcdA. We used an SspB-obligate substrate (GFP-ssrA-SS) that requires an adaptor for ClpXP degradation (Lau et al., 2015) (Figure 3F). Consistent with our model that the RcdA C-terminal tether relies on a CpdR-primed ClpXP protease, the SspB~RcdA chimeric adaptor fails to deliver GFP-ssrA-SS to ClpXP alone but robustly promotes substrate degradation when CpdR is added (Figure 3F).

Next, we directly measured RcdA binding to its protease partner by following anisotropy of a fluorescently labeled peptide consisting of the last 19 residues of RcdA. This peptide bound poorly to either CpdR or ClpX alone, but addition of both resulted in strong peptide binding (Figure 3G). Excess RcdA competitively inhibited peptide binding, but RcdAΔC could not compete, consistent with the above results illustrating that the C terminus of RcdA is the major contact with the CpdR-primed ClpX (Figures 3G and S2D). CpdR binds the unique N-terminal domain of ClpX as part of its delivery mechanism (Lau et al., 2015), and we found that RcdA peptide binding and substrate delivery to the CpdR-ClpX complex also require this region (Figures 3G and S2E). Taken together, these results support a model in which the CpdR adaptor binds to the N-terminal domain of ClpX, priming it to recognize the C-terminal disordered region of RcdA and enhance degradation of the TacA substrate.

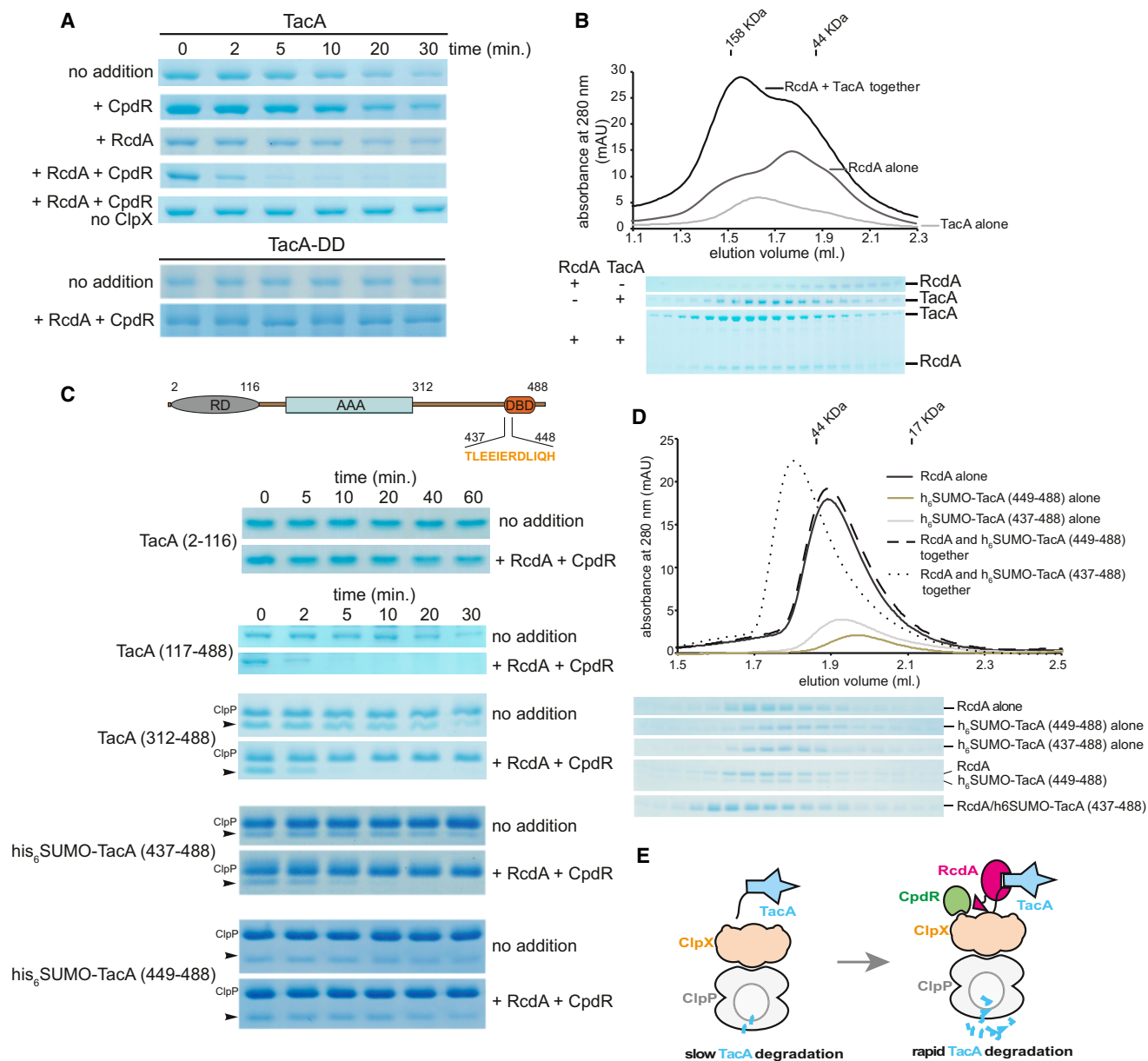


Figure 2. RcdA Binds the TacA Substrate Directly to Promote Degradation by ClpXP

(A) RcdA and CpdR collaborate to accelerate TacA degradation in vitro. Degradation reactions were performed using either RcdA or CpdR alone or together. Reaction containing both RcdA and CpdR but not ClpX served as a control. Reactions consisted of 1 μ M TacA, 1 μ M TacA-DD, 1 μ M RcdA, 2 μ M CpdR, 0.4 μ M ClpX₆, and 0.8 μ M ClpP₁₄ when indicated. See also Figures S1A and S1B.

(B) RcdA directly binds to TacA. His₆-TacA or His₆-RcdA were loaded either alone or together onto the analytical size-exclusion column. Collected fractions were resolved by SDS-PAGE gels. Protein standards used to calibrate the column are indicated above.

(C) C-terminal region of TacA contains the necessary recognition element for RcdA/CpdR-mediated degradation. Both N- and C-terminally truncated variants of TacA as indicated by residue numbers were generated and subjected to RcdA/CpdR-mediated ClpXP degradation. Degradation reaction consisted of 8 μ M TacA (2-116), 1 μ M TacA (117-488), 3 μ M TacA (312-488), 2 μ M his₆SUMO-TacA (437-488), 2 μ M his₆SUMO-TacA (449-488), 1 μ M RcdA, 2 μ M CpdR, 0.2 μ M ClpX₆, and 0.4 μ M ClpP₁₄ when indicated. Arrows indicate bands corresponding to the TacA variants.

(D) RcdA binds the C terminus domain of TacA containing residues 437-488 but not a domain lacking residues 437-448. RcdA- or SUMO-appended TacA (437-488) or TacA (449-488) variants were loaded either alone or together onto the analytical size-exclusion column. Collected fractions were resolved by SDS-PAGE gels and stained by Coomassie. Protein standards used to calibrate the column are indicated above. See also Figure S1D.

(E) Cartoon of working model wherein RcdA binds TacA and delivers it to a CpdR-bound ClpXP.

Tethering of the RcdA Adaptor to ClpX Bypasses the Need for CpdR

Our results suggested that tethering the RcdA adaptor to a CpdR-primed ClpXP complex is needed for TacA degradation. If this model is correct, then we should be able to constitutively activate RcdA as an adaptor by directly tethering it to ClpX. We fused the ClpX-binding motif of SspB to the RcdAΔC variant to generate the RcdAΔC~XB chimera. In accordance with our model, RcdAΔC~XB was able to stimulate the degradation of TacA without the need for CpdR (Figure 4A). Strains lacking CpdR have higher steady-state levels of TacA than WT strains, as there is a loss of CpdR-dependent TacA degradation (Figures 4B and 4C). Importantly, expression of RcdAΔC~XB in this background reduces TacA to WT levels (Figure 4B) due to a recovery of TacA degradation in a CpdR-independent manner (Figure 4C).

RcdA is known to interact with PopA, a cyclic di-GMP (cdG) binding protein (Duerig et al., 2009). We recently showed that RcdA, PopA, and CpdR form a multi-protein adaptor complex in the presence of cdG that delivers CtrA for degradation by ClpXP (Duerig et al., 2009; Smith et al., 2014). The ability of RcdA to aid in this adaptor complex formation requires CpdR (Smith et al., 2014), and we speculated that this extended function of RcdA could also be bypassed by RcdAΔC~XB (Figure 4D). Consistent with this hypothesis, adding RcdAΔC~XB, PopA, and cdG stimulated the degradation of CtrA even without CpdR (Figures 4E and S3). Similar to the recovery of TacA degradation shown above, degradation of CtrA in vivo could also be restored in cells lacking CpdR if RcdAΔC~XB was expressed (Figure 4F). These results demonstrate that direct tethering of RcdA to ClpX is sufficient for delivery of substrates to ClpXP and to assemble a multi-protein adaptor complex for CtrA degradation. Thus, priming of ClpX by the CpdR adaptor contributes to CtrA/TacA degradation principally by generating a tethering site for the RcdA adaptor.

Proteomic Identification of Additional RcdA-Dependent ClpXP Substrates

Our working model is that the RcdA adaptor recognizes protease substrates via an N-terminal substrate-binding domain and engages a CpdR-primed ClpX via its disordered C-terminal region. We speculated that other substrates in addition to TacA may also be delivered to ClpXP in a similar fashion and used a proteomics approach to identify these candidates. We expressed epitope-tagged RcdA variants M2-RcdA (which fully complements; Figure S4A) and M2-RcdAΔC in ΔrcdA cells, lysed cells, and precipitated RcdA-interacting proteins using M2-FLAG affinity beads (Figure 5A). Our rationale for using both variants was that targets of RcdA should bind both constructs but would be enriched in the M2-RcdAΔC binding pool due to the cell's inability to degrade those targets. As expected, TacA was enriched in the elution fraction of M2-RcdAΔC relative to the M2-RcdA elution and was absent in mock pull-down experiments (Figure S4B). We identified putative RcdA partners by trypsinization/mass spectrometry and applied an enrichment filter to prune the candidate pool (see Experimental Procedures and Supplemental Experimental Procedures). We focused on two proteins of unknown function that were strongly enriched by this approach

(see Experimental Procedures and Supplemental Experimental Procedures). CC2323 was a protein that we had previously identified as a ClpXP substrate based on a ClpP trapping approach (Bhat et al., 2013), and CC3144 was a protein that was enriched as strongly as TacA in our RcdA pull down (see Experimental Procedures and Supplemental Experimental Procedures).

We found that CC3144 protein levels oscillated in a cell-cycle-dependent manner similar to TacA and CtrA (Figure 5B). Importantly, purified CC3144 was only robustly degraded upon addition of CpdR and RcdA (Figure 5C). CpdR and RcdA were both required for CC3144 degradation in vivo, but as in the case for TacA, PopA was dispensable for CC3144 degradation (Figure 5B). Similarly, RcdA forms a tight complex with purified CC2323 in vitro (Figure S4C), and degradation of this protein by ClpXP was enhanced by addition of CpdR and RcdA (Figure 5D), although an M2-tagged variant of CC2323 did not mirror TacA or CC3144 degradation in vivo, possibly due to disruption of the protein by the tag (Figure S4D). Taken together, these data reveal that the RcdA adaptor can bind and deliver a number of substrates in addition to TacA (Figure 5E).

Adaptors Can Also Act as Anti-adaptors

Adaptors can be controlled by anti-adaptors, as shown by the complex regulation of the RssB adaptor by the Ira family of proteins (Bougoudour et al., 2006; Battesti et al., 2013b). Outside of RssB, few examples of anti-adaptors have been described; however, we considered that RcdA could be subject to anti-adaptor regulation as it can clearly bind diverse partners (Figure 5). In particular, PopA must be bound to cdG in order to facilitate degradation of CtrA by CpdR/RcdA/ClpXP (Figure S3) (Smith et al., 2014), but prior work suggests that PopA and RcdA still bind in the absence of cdG (Duerig et al., 2009; Smith et al., 2014). Therefore, we speculated that PopA could act as an inhibitory anti-adaptor for RcdA-specific substrates in addition to its stimulatory role in activating CtrA degradation (Figure 6A). Indeed, steady-state levels of TacA and CC3144 are slightly lower in ΔpopA cells (Figures 6B and S5A). Consistent with an anti-adaptor role for PopA, TacA is degraded more rapidly in ΔpopA cells (Figures 6C, S5B, and S5C). Excess PopA suppresses RcdA-dependent TacA degradation in vitro even in the absence of cdG (Figure 6D), whereas addition of cdG promotes CtrA degradation (Figure 6E). Thus, PopA can be both an adaptor of RcdA that enhances CtrA degradation and an anti-adaptor of RcdA that blocks TacA degradation. Similarly, excess RcdA inhibits degradation of the PdeA substrate in vitro, and steady-state levels of the McpA substrate are lower in the absence of RcdA and PopA (Figures S5D and S5E), consistent with a general model wherein members of the adaptor hierarchy can act as anti-adaptors for substrates at a lower hierarchical level.

DISCUSSION

Protein degradation is essential but must be exquisitely controlled so that off-target proteins are not destroyed, as unrestrained proteolysis is lethal to cells (Brötz-Oesterhelt et al., 2005). In *Caulobacter crescentus*, the regulated destruction of cell-cycle factors by the essential protease ClpXP coordinates

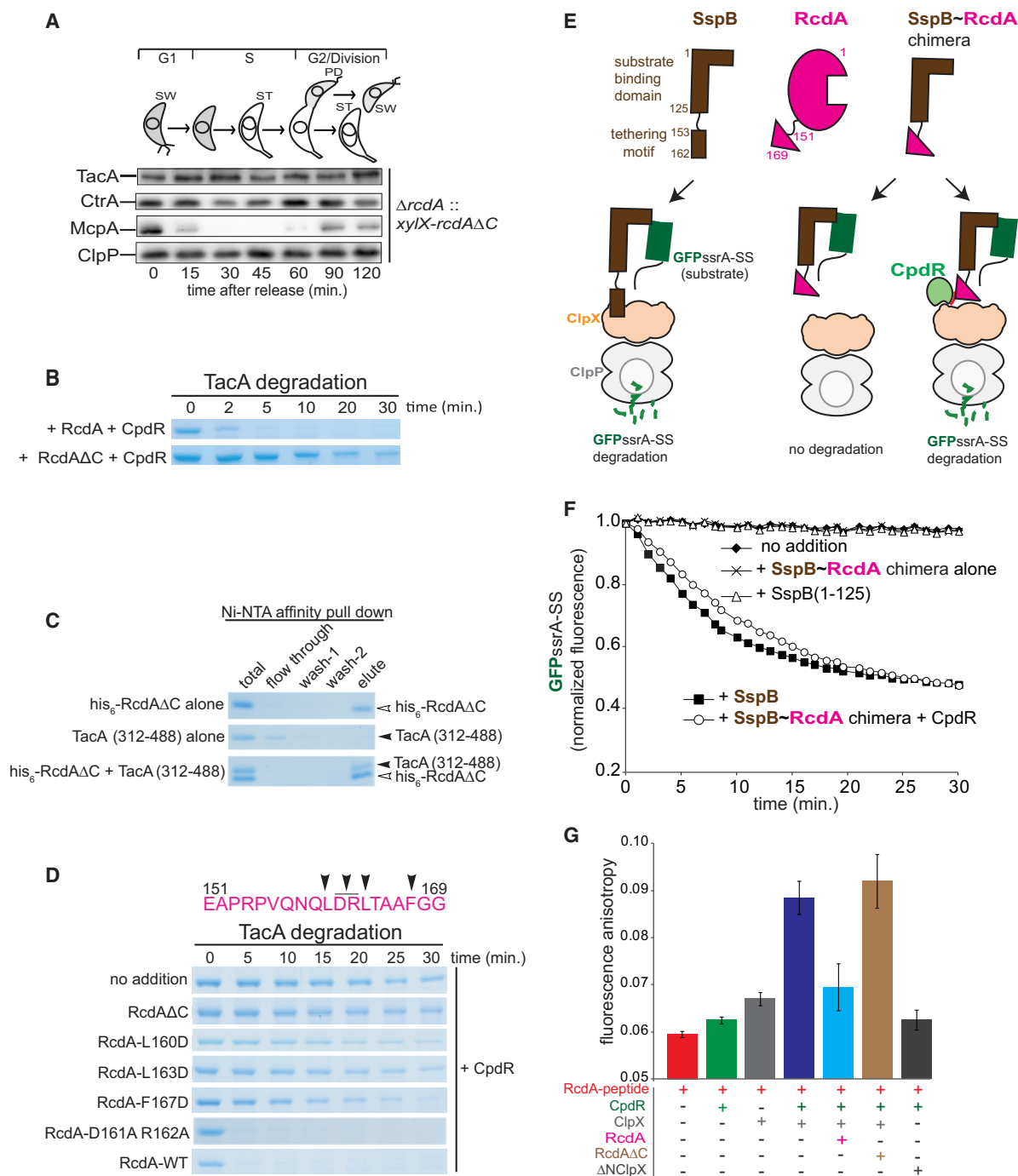


Figure 3. RcdA Contains a Tethering Motif Needed for Substrate Delivery to the CpdR-Primed ClpXP

(A) Cell-cycle-dependent degradation of TacA requires the C-terminal region of RcdA. Strain expressing RcdAΔC (RcdA lacking the 19-residue C-terminal region) was synchronized, and isolated swarmer cells were released into fresh PYE medium. Constant volume of samples was withdrawn at indicated time points for western blot analysis. Each sample was probed with anti-TacA, anti-CtrA, anti-McpA and anti-ClpP antibodies.

(B) TacA degradation is not stimulated by RcdAΔC in vitro even in the presence of CpdR. Degradation reactions were performed with either purified full-length RcdA or RcdAΔC in the presence of CpdR. Degradation reactions were similar to those in Figure 1A except that 1 μM RcdAΔC was used.

(C) RcdAΔC directly binds the C-terminal domain of TacA. His₆-RcdAΔC was used as a bait protein to pull down TacA (312–488) by using Ni-NTA affinity resin. See also Figures S2A and S2B.

(D) Hydrophobic residues in the C-terminal region of RcdA are important for RcdA-mediated delivery of TacA. Residues L160, L163, and F167 were mutated to aspartate, and residues D161 and R162 were mutated to alanines (arrows). RcdAΔ161A,R162A was used as a control. See also Figure S2C. TacA degradation was monitored in reactions containing these purified RcdA mutants and CpdR.

(legend continued on next page)

replication and development. Genetic studies pointed to several key factors necessary to mediate degradation *in vivo*, but it was unclear how they mechanistically coordinated regulated degradation. Part of this uncertainty arose from the complex relationship between these factors and substrate degradation, wherein different factors were needed for degradation of different substrates. Although ClpXP is active on its own, the response regulator CpdR is required for degradation of all known cell-cycle-dependent ClpXP substrates *in vivo* (Iniesta et al., 2006; Abel et al., 2011; Bhat et al., 2013), and the proteins RcdA and PopA are needed for degradation of only a subset of these targets (Duerig et al., 2009; Radhakrishnan et al., 2010). We found previously that the adaptor CpdR directly facilitates degradation of some ClpXP substrates, such as the phosphodiesterase PdeA (Rood et al., 2012; Lau et al., 2015); however, the CpdR adaptor alone is insufficient to enhance degradation of other substrates such as CtrA (Smith et al., 2014). Here, we show that RcdA is an adaptor that tethers proteins to the ClpXP protease only when the protease is first primed by CpdR. The RcdA adaptor binds the protease substrates TacA, CC2323, and CC3144 to deliver them to a CpdR-primed ClpXP. RcdA also binds the PopA adaptor, which promotes recruitment of additional substrates, such as CtrA, to the primed protease (Figure 7).

Implications for Adaptor and Substrate Discovery

Traditionally, an adaptor protein could be recognized by its ability to bind substrate directly and to physically interact with the protease to deliver the substrate. Examples of ClpXP adaptors include SspB, which binds and delivers ssrA-tagged substrates, and RssB, which delivers the RpoS substrate to ClpXP (Zhou et al., 2001; Dougan et al., 2003). For both cases, the definitive proof that these were adaptors was that they were able to enhance degradation of their substrates by ClpXP *in vitro*. It is inherently difficult to identify new adaptors because without prior knowledge of their substrates, designing an experiment to test adaptor activity is almost impossible. For example, in this work we found that RcdA was able to deliver substrates for ClpXP-mediated degradation. However, RcdA cannot deliver substrates to ClpXP alone but requires a CpdR-primed ClpXP. Had we not initially identified CpdR as an adaptor of ClpXP during our characterization of PdeA degradation (Abel et al., 2011; Lau et al., 2015), our efforts to reconstitute the adaptor activity of RcdA would have been futile. This example demonstrates a central difficulty in adaptor/substrate discovery. Namely, a candidate is proven to be an adaptor by monitoring the degradation of a specific substrate, but reconstituting substrate degradation requires knowing that the adaptor or other additional factors are needed in the first place. Systematic

approaches to identifying new adaptor/substrate pairs are needed.

Proteolytic Regulation by Hierarchical Assembly of Adaptors

Protein degradation is a tightly regulated process that is needed for cell-cycle progression. In eukaryotes, the 26S proteasome specifically recognizes ubiquitinated proteins generated by the stage-specific activity of APC/C and SCF ubiquitin ligases. As bacteria lack ubiquitination, they often employ adaptors to promote rapid protein degradation. How this process can accommodate the hundreds of substrates degraded during the cell cycle (Duerig et al., 2009; Radhakrishnan et al., 2010; Bhat et al., 2013; Smith et al., 2014) while maintaining specificity is a substantial challenge. One way bacteria can overcome this challenge is to employ hierarchical control of proteases, tightly regulating the activity of a single protein during the cell cycle that initiates a cascade of proteolytic events. In the case of *Caulobacter crescentus*, cell-cycle-dependent phosphorylation of CpdR gates its ability to prime ClpX for recruitment of substrates such as PdeA (Abel et al., 2011; Lau et al., 2015). When activated, CpdR binding to ClpX also promotes engagement of the RcdA adaptor that directly delivers some substrates (TacA, CC2323, CC3144) and indirectly delivers substrates such as CtrA through recruitment of additional adaptors such as PopA. Thus, which substrate is degraded depends on the degree of the adaptor hierarchy that has assembled at that time, which could aid in prioritization of substrate destruction, e.g., CtrA is only degraded when the complete adaptor hierarchy is assembled (Figure 7).

When bound to cdG, PopA acts as an adaptor to promote CtrA degradation, but in the absence of cdG, PopA can moonlight as an anti-adaptor for RcdA-dependent protease substrates such as TacA. Similarly, excess RcdA adaptor can inhibit degradation of PdeA presumably by competing for a limited amount of CpdR-primed ClpXP. We speculate that a general feature of adaptor hierarchies is that adaptors operating at a higher level of the hierarchy can serve as anti-adaptors for substrates reliant only on the lower levels of the hierarchy. An intriguing corollary to this model is that other known anti-adaptors (such as the Ira family of proteins that block the RssB adaptor) could also moonlight as adaptors and aid in the delivery of as yet unknown substrates.

The Role of Adaptor Hierarchies in Bacterial Development

Finally, we again note that CpdR and RcdA are conserved throughout α -proteobacteria, whereas the presence of PopA appears restricted to *Caulobacter* and closely related species

(E) Cartoon depicting SspB, RcdA, and SspB~RcdA chimeric proteins and how delivery of GFP-ssrA-SS substrate is affected by adaptors. Residue numbers indicate boundaries used to construct the chimeric protein.

(F) SspB~RcdA chimera delivers GFP-ssrA-SS to ClpXP in a CpdR-dependent manner. GFP-ssrA-SS was subjected to SspB~RcdA chimera-mediated degradation in the presence or absence of CpdR. Loss of GFP fluorescence was monitored over time. Reaction consisted of 2 μ M GFP-ssrA-SS, 0.5 μ M SspB, 0.5 μ M SspB (1–125), 1 μ M SspB-RcdA chimera, 4 μ M CpdR, 0.4 μ M ClpX₆, and 0.8 μ M ClpP₁₄ when indicated.

(G) The isolated RcdA C-terminal peptide binds to CpdR-primed ClpX but not ClpX variant lacking the N-terminal domain. Fluorescence anisotropy was measured for the binding of 40 nM FITC-labeled, C-terminal 19-residue peptide of RcdA to 4 μ M CpdR, 1 μ M ClpX₆ or 1 μ M Δ NClpX₆, 10 μ M RcdA, and 10 μ M RcdA Δ C, either alone or in different combinations as indicated.

Data represent mean \pm SD. See also Figures S2D and S2E.

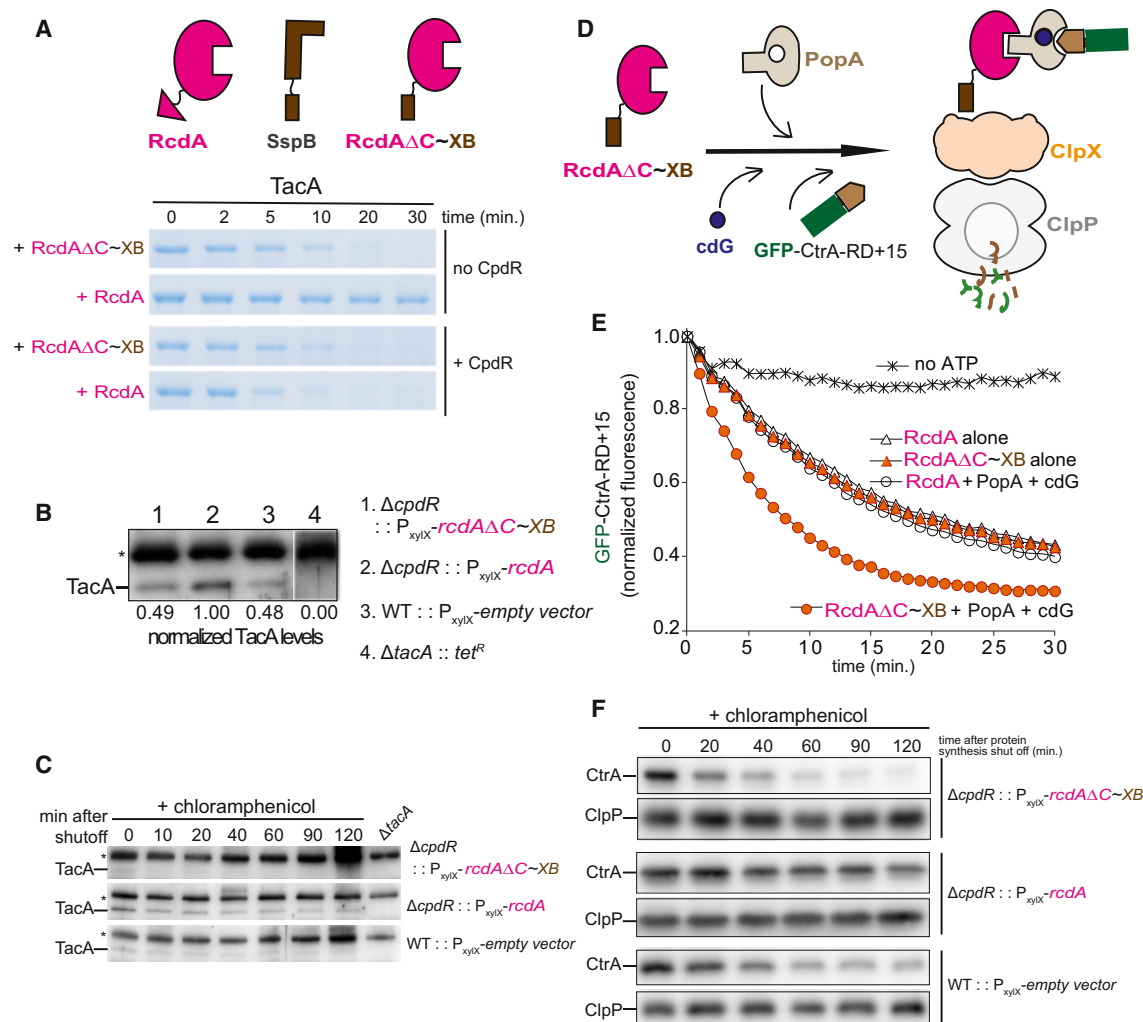


Figure 4. Direct Tethering of RcdA to ClpX Is Sufficient for Delivery of the TacA Substrate

(A) RcdAΔC~XB alone delivers TacA to ClpXP for degradation in vitro. Schematic depicting different fusion constructs is shown. Degradation reactions were performed with RcdAΔC~XB alone or with RcdA alone or in the presence of CpdR. Degradation reactions consisted of 1 μM TacA, 1 μM RcdAΔC~XB, 1 μM RcdA, 2 μM CpdR, 0.4 μM ClpX₆, and 0.8 μM ClpP₁₄ when indicated.

(B and C) RcdAΔC~XB delivers TacA for degradation in the absence of CpdR in vivo. Steady-state levels of TacA were measured in Δ*cpdR* cells expressing *rcdAΔC~XB*, *rcdA* or WT cells containing the empty vector. Exponential phase cells were induced with 0.003% xylose for 1 hr. Lysate from equal numbers of cells was used for western blot analysis to probe using anti-TacA antibody. Cropped image of the blot from Δ*tacA* lysate is shown. Protein synthesis was inhibited during exponential growth with chloramphenicol. Samples were withdrawn at indicated time points and normalized to OD before western blot analysis. Asterisk indicates a cross-reacting band.

(D and E) RcdAΔC~XB in conjunction with PopA and cdG delivers a GFP-CtrA reporter to ClpXP independent of CpdR in vitro. Cartoon depicting RcdAΔC~XB/PopA/cdG-mediated GFP-CtrA-RD+15 delivery to ClpXP is shown. Reactions for monitoring loss of fluorescence from GFP-CtrA-RD+15 degradation were carried out with RcdA alone, RcdA with PopA/cdG, RcdAΔC~XB alone, or RcdAΔC~XB with PopA/cdG. Degradation reactions consisted of 1 μM GFP-CtrA-RD+15, 1 μM RcdAΔC~XB, 1 μM RcdA, 2 μM CpdR, 1 μM PopA, 20 μM cdG, 0.4 μM ClpX₆, and 0.8 μM ClpP₁₄ when indicated. See also Figure S3.

(F) RcdAΔC~XB promotes CtrA degradation in the absence of CpdR in vivo. Experiments were performed as in panel C, except that blots were probed with anti-CtrA and anti-ClpP antibodies. ClpP was used as a loading control.

(Brilli et al., 2010; Ozaki et al., 2014). As PopA is responsive to the second messenger cyclic di-GMP, levels of which oscillate during the cell cycle (Abel et al., 2013; Lori et al., 2015), our work suggests that CpdR and RcdA may represent an ancestral adaptor complex that has been co-opted by *Caulobacter* for coupling second messenger cues to cell-cycle progression. In *S. meliloti*, it is clear that adaptors play a crucial role in the symbiosis transition, wherein misregulation of CpdR dramatically

affects proper nodule formation and plant growth (Kobayashi et al., 2009; Pini et al., 2015; Schallies et al., 2015). In general, bacterial development (such as the morphological transition of *Caulobacter* or sporulation of *Bacillus*) requires changes in proteome composition within a single generation. Adaptor hierarchies could robustly promote these changes while maintaining the control needed to selectively degrade proteins in a specific order or priority.

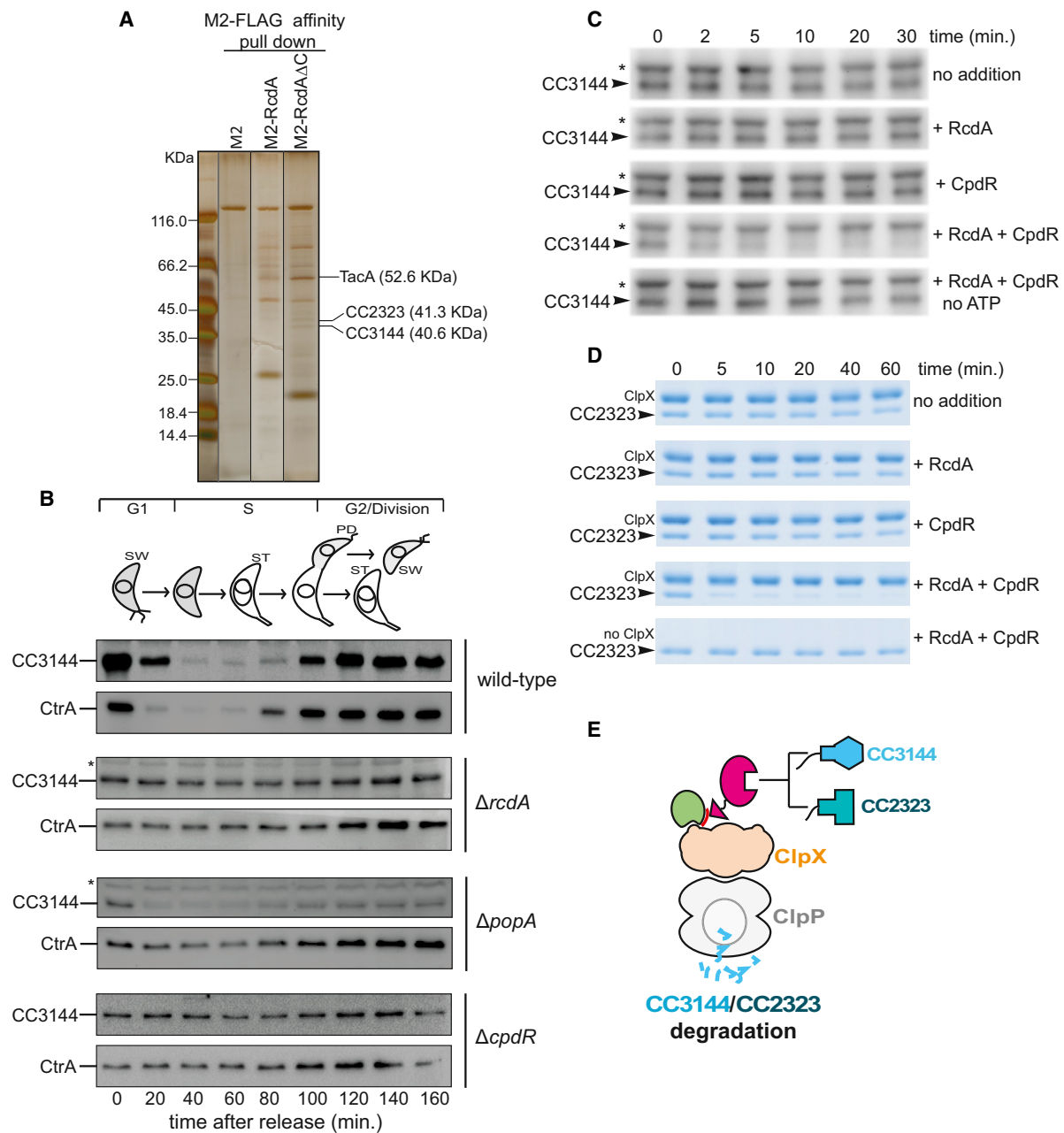


Figure 5. Proteomic Identification and Validation of RcdA-Dependent ClpXP Substrates

(A) In vivo pull downs reveal additional binding partners of RcdA. M2-FLAG affinity resin was used to pull down M2-epitope tagged RcdA or RcdA Δ C from lysates of $\Delta rcdA$ cells expressing either M2-RcdA or M2-RcdA Δ C. Lysate from cells expressing M2 peptide alone was used as a control. Cropped images of silver-stained SDS-PAGE gels from elution pool of M2, M2-RcdA, and M2-RcdA Δ C are shown. Protein markers are indicated. See also Figures S4A and S4B.

(B) CC3144 is degraded in a cell-cycle-dependent manner requiring RcdA and CpdR but not PopA. Cultures of WT or cells lacking RcdA, CpdR, and PopA were synchronized, and swarmer populations were released into fresh M2G medium. Constant volume of samples were withdrawn at indicated time points and probed with anti-CC3144 and anti-CtrA antibodies.

(C and D) ClpXP-mediated degradation of CC3144 or CC2323 is enhanced in the presence of RcdA and CpdR. Reactions consisted of 1 μ M CC3144 or CC2323, 1 μ M RcdA, 2 μ M CpdR, 0.4 μ M ClpX₆, and 0.8 μ M ClpP₁₄ when indicated. In (D), CC2323 degradation was visualized by Coomassie staining. In (C), overlapping bands made it necessary to use western blotting to detect the purified CC3144 using anti-CC3144 antibody. Asterisks denote cross-reacting bands. See also Figure S4C.

(E) Cartoon of RcdA delivering diverse substrates to a CpdR-primed ClpXP.

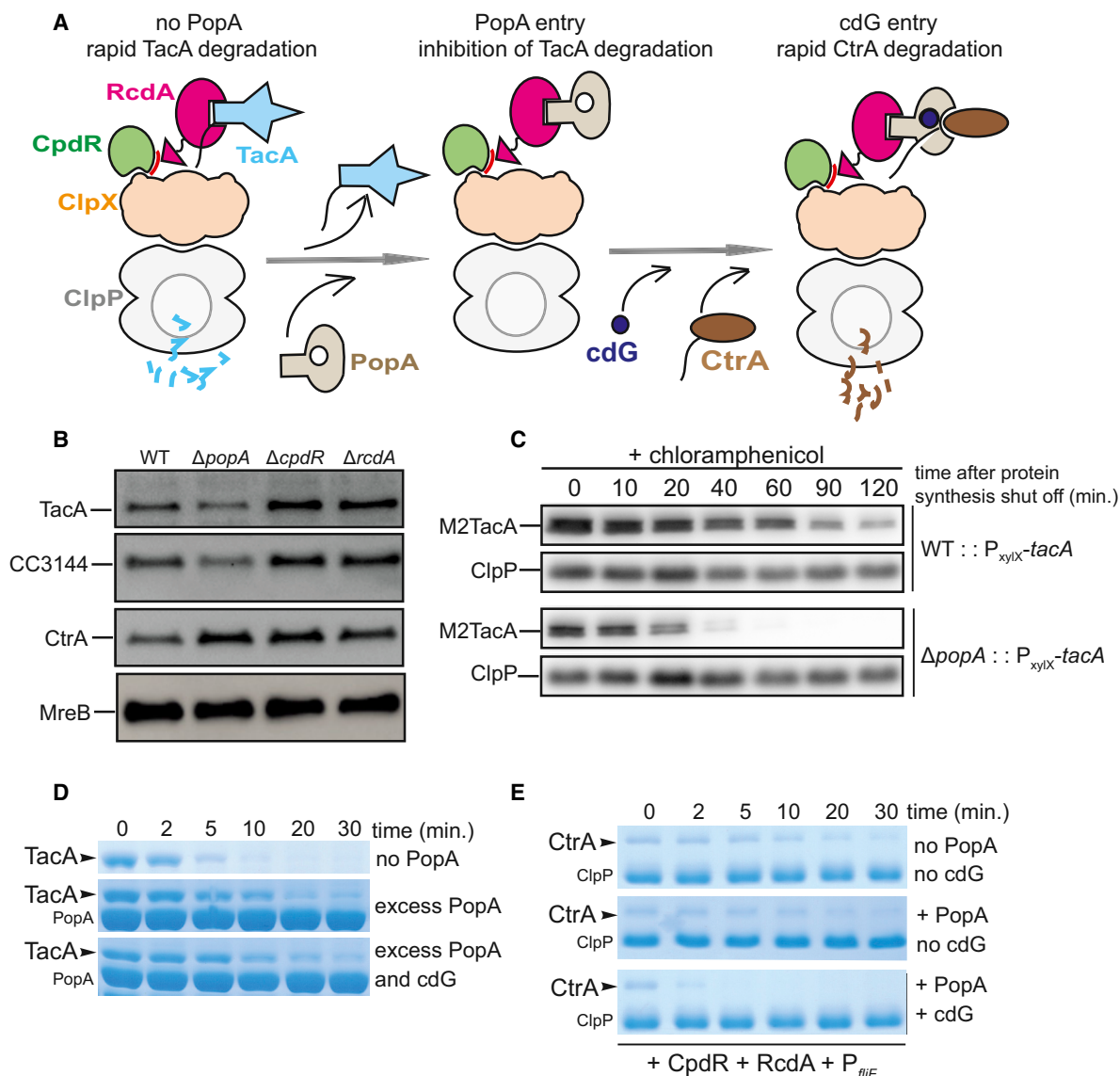


Figure 6. Adaptors Can Also Act as Anti-adaptors

(A) Cartoon depicting PopA-mediated inhibition of TacA degradation and PopA/cdG-dependent stimulation of CtrA degradation.

(B) Steady-state levels of TacA and CC3144 are lower in $\Delta popA$ cells compared to WT, $\Delta rcdA$, or $\Delta cpdR$ cells. Lysates from equal OD₆₀₀ of WT, $\Delta popA$, $\Delta rcdA$, and $\Delta cpdR$ strains were used for western blot analysis, probing with anti-TacA, anti-CC3144, anti-CtrA, and anti-MreB antibodies. MreB was used as a loading control. See also Figure S5A.

(C) TacA is degraded more rapidly in $\Delta popA$ cells compared to WT. WT and $\Delta popA$ cells bearing M2-epitope-tagged TacA were grown at exponential phase and subsequently induced with 0.003% xylose for 1 hr. Protein synthesis was inhibited by addition of chloramphenicol. Samples were withdrawn at indicated time points and normalized to OD before performing western blot analysis using anti-TacA and anti-ClpP antibodies. ClpP was used as a loading control. See also Figures S5B and S5C.

(D) Excess PopA suppresses TacA degradation independent of cdG in vitro. Reactions consist of 1 μ M TacA, 1 μ M RcdA, 2 μ M CpdR, 5 μ M PopA, 100 μ M cdG, 0.4 μ M ClpX₆, and 0.8 μ M ClpP₁₄ when indicated.

(E) PopA/CpdR/RcdA accelerates CtrA degradation in a cdG-dependent manner in vitro. Reaction conditions are similar to those in Figure 6D, except that 1 μ M PopA, 20 μ M cdG, and 10 μ M P_{fliF} were used.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions

Bacterial strains used in the study are tabulated in Table S2. *E. coli* strains were grown in Luria-Broth (LB) at 37°C with the appropriate antibiotic

(100 μ g/ml ampicillin, 50 μ g/ml kanamycin, 50 μ g/ml spectinomycin). *Caulobacter* strains were grown in Peptone-Yeast -Extract (PYE) media at 30°C with 25 μ g/ml spectinomycin or 5 μ g/ml kanamycin, wherever required and supplemented with xylose to induce gene expression as indicated in the figure legends.

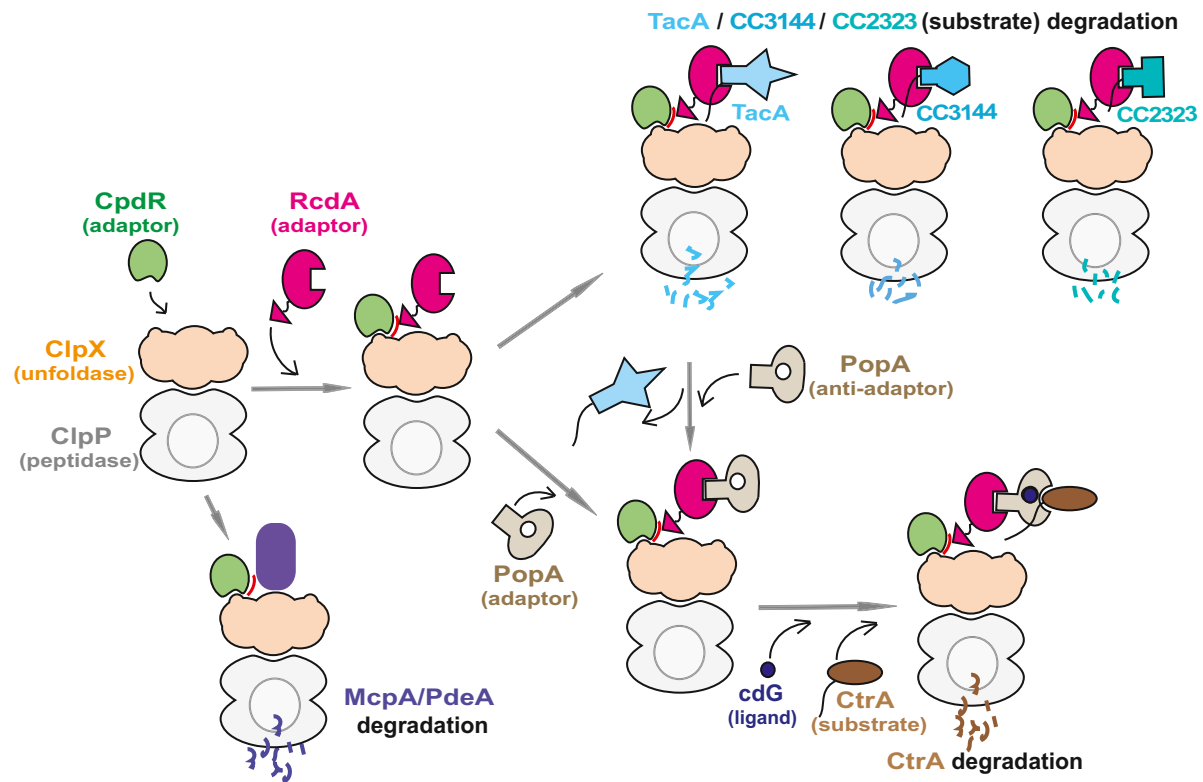


Figure 7. A Model for Adaptor Hierarchies Orchestrating Proteolysis

Priming of the ClpXP protease by the CpdR adaptor enhances degradation of the substrates McpA and PdeA. The CpdR-primed ClpXP complex can also recruit the RcdA adaptor. RcdA binds directly to the substrates TacA, CC3144, or CC2323 and delivers them to the CpdR-primed ClpXP protease for degradation. The PopA protein acts as an anti-adaptor by binding the RcdA adaptor and inhibiting degradation of the substrate TacA. In response to cdG, PopA also acts as an adaptor of RcdA to deliver the substrate CtrA to the CpdR-primed ClpXP protease. This hierarchical formation of adaptors, anti-adaptors, and primed proteases dictates cell-cycle-dependent protein degradation.

Cloning, Mutagenesis, and Protein Purification

Caulobacter strains expressing TacA and TacA-DD were used as described previously (Bhat et al., 2013). *Caulobacter* strain harboring the *rcdAΔC* allele was used as described by Taylor et al. (2009). M2-tagged RcdA or M2-tagged RcdAΔC was generated in pENTR plasmid and transferred into xylose-inducible expression plasmids using Gateway-based cloning (Skerker et al., 2005). Truncated variants of TacA were generated by amplifying the desired region of TacA, as indicated in the figures, using round-the-horn site-directed mutagenesis or Gibson assembly method (Gibson et al., 2009) with pET23SUMO as template. CC3144 and CC2323 were amplified with appropriate primers using genomic DNA from *C. crescentus* as template and then cloned into pTE28a and pET23SUMO expression plasmid, respectively. SspB-RcdA chimeric protein was generated by replacing 10 C-terminal residues of SspB with 19 C-terminal residues of RcdA. The RcdAΔC~XB fusion construct was generated in pET28a vector by designing appropriate primers to append the last 10 residues of SspB to RcdAΔC. All the constructs were confirmed by sequencing. Oligonucleotide sequences are available upon request.

BL21(DE3) *E. coli* cells bearing expression plasmid for different proteins were grown till the optical density 600 (OD₆₀₀) reached ~0.4 to 0.8, induced with 0.4 mM IPTG for 3–5 hr and then centrifuged at 5,000 rpm for 15 min. Pellets were resuspended in lysis buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl, 10 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol and frozen at –80°C until further use. Cells were lysed using a Microfluidizer system (Microfluidics, USA). The lysate was applied onto a Ni-NTA column for affinity purification. SUMO-tagged proteins were cleaved by Ulp1-his protease (Rood et al., 2012). Proteins were further purified by size-exclusion

and anion-exchange chromatography using Sephacryl 200 16/60 and MonoQ 5/50 columns. ClpX and ClpP were purified as described previously (Chien et al., 2007). Detailed purification protocols are available upon request.

Synchronization and In Vivo Protein Stability Assays

Synchronization experiments were performed by growing *Caulobacter* strains to an OD₆₀₀ of 0.3–0.5 in PYE or M2G media with appropriate antibiotic and xylose when required, as indicated in the figure legends. Swarmer cells were isolated using Percoll density gradient centrifugation. The isolated swarmer cells were then released into fresh PYE or M2G minimal media for progression through the cell cycle. In vivo protein stability assays were performed by growing WT or *Caulobacter* cells, expressing different constructs from a xylose-inducible plasmid, to an OD₆₀₀ of ~0.3 in PYE medium containing 25 μg/ml spectinomycin. Protein expression was induced with 0.003% xylose for 1 hr wherever indicated in figure legends, and then protein synthesis was blocked by addition of 30 μg/ml chloramphenicol.

Microscopy

Phase-contrast microscopy was performed on glass slides layered with a 1% agarose pad. A Zeiss Scope A.1 microscope (Zeiss, Germany) equipped with 100× (1 × 25 oil ph3 ∞/0.17) objective and 60 N-C 1" 100× camera was used. Images were analyzed with Axiovision and ImageJ (NIH, USA) software.

Gel and Fluorescence-Based Degradation Assays

Degradation of proteins was monitored using SDS-PAGE gels as described previously (Bhat et al., 2013). The concentrations of different proteins used in degradation reactions are indicated in the figure legends. Degradation of

GFP-ssrA and GFP-ssrA-SS were monitored with the loss of fluorescence over time as described previously (Smith et al., 2014).

Western Blot Analysis

Cell cultures withdrawn at indicated time points were spun down, resuspended in 2× SDS sample buffer, boiled at 95°C for 10 min, and then centrifuged. After centrifugation, the clarified supernatant was loaded onto SDS-PAGE gels. Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane at 20V for 1 hr and probed with polyclonal rabbit anti-CtrA (1:5000 dilution), anti-McpA (1:10000 dilution), anti-TacA (1:10000 dilution), anti-CC3144 (1:5000 dilution), anti-ClpP (1:5000 dilution), anti-MreB (1:10000 dilution), or monoclonal mouse anti-FLAG M2 (1:5000 dilution; Sigma, USA) antibodies. Proteins were visualized with either goat anti-rabbit (Millipore, USA) or goat anti-mouse (Millipore, USA) secondary antibodies conjugated to HRP using a chemiluminescence detection system (G:box Chemi XT4, Syngene, UK).

Analytical Size-Exclusion Chromatography

Analytical size-exclusion chromatography was performed to detect interaction between proteins using a Superdex 200 5/150 column (GE Healthcare, USA). The column was equilibrated with H-buffer (20 mM HEPES-KOH, pH 7.5, 100 mM KCl, 10 mM MgCl₂, 10% glycerol, 5 mM β-mercaptoethanol) and standardized using thyroglobulin, γ-globulin, ovalbumin, myoglobin, and vitaminB₁₂ as protein standards (Biorad, USA). Purified RcdA or RcdAΔC either alone or in combination with TacA or TacA variants was loaded onto the column. The elution profile was monitored by measuring absorbance at 280 nm.

In Vitro and In Vivo Pull-Down Assays

In vitro pull-down assays were performed by incubating purified his₆-RcdA (10 μM) or his₆-RcdAΔC (10 μM) or TacA (312–488) (5 μM) or PdeA (5 μM) either alone or together with 50 μl pre-equilibrated Ni-NTA resin (Thermo-Scientific, USA) at 4°C for 1 hr. The resin was washed twice with H-buffer supplemented with 20 mM imidazole. Bound complex was eluted with H-buffer containing 200 mM imidazole. Eluted proteins were analyzed by SDS-PAGE gels.

In vivo pull-down assays were performed by loading lysates (containing equal amount of protein) from cells expressing either M2 peptide, M2-RcdA, or M2-RcdAΔC onto a pre-equilibrated anti-FLAG M2 affinity resin packed column (Sigma, USA). The buffer used for equilibration was 50 mM Tris (pH 7.5) and 150 mM NaCl (TBS). The resin was washed twice with TBS, and the bound proteins were eluted with 0.1 M glycine-HCl (pH 3.5) buffer. Eluted proteins were analyzed by silver staining and identified by tandem mass spectrometry (see Supplemental Experimental Procedures).

Fluorescence Anisotropy-Based Assay

Fluorescence anisotropy experiments were performed using a fluorescein isothiocyanate (FITC) labeled, 19-residue C-terminal peptide of RcdA (FITC-EAPRPVQNQLDRLTAAGG, LifeTein, USA). As indicated in Figure 3G, the peptide was incubated with proteins in H-buffer. Anisotropy was monitored at 30°C with a Spectramax M5 (Molecular Devices) plate reader with excitation and emission wavelengths set at 488 nm and 525 nm, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.030>.

AUTHOR CONTRIBUTIONS

Conceptualization, K.K.J. and P.C.; investigation, K.K.J., M.B., and S.K.R.; writing—original draft, K.K.J. and P.C.; writing—review and editing, K.K.J., M.B., P.H.V., and P.C.; funding acquisition, P.C. and P.H.V.; supervision, P.C. and P.H.V.

ACKNOWLEDGMENTS

We thank the Chien and Viollier lab members, E. Vierling, D. Hebert, and C. Babbitt for helpful comments and discussions on the manuscript. We also thank Kathleen R. Ryan (UC Berkeley) for providing us plasmids and strains. K.K.J. thanks Amrita Palaria for helpful discussions. This project was supported by funds from University of Massachusetts Amherst and NIH grant R01GM111706 (to P.C.) and SNF #31003A_143660 (to P.H.V.).

Received: May 15, 2015

Revised: July 1, 2015

Accepted: September 2, 2015

Published: October 8, 2015

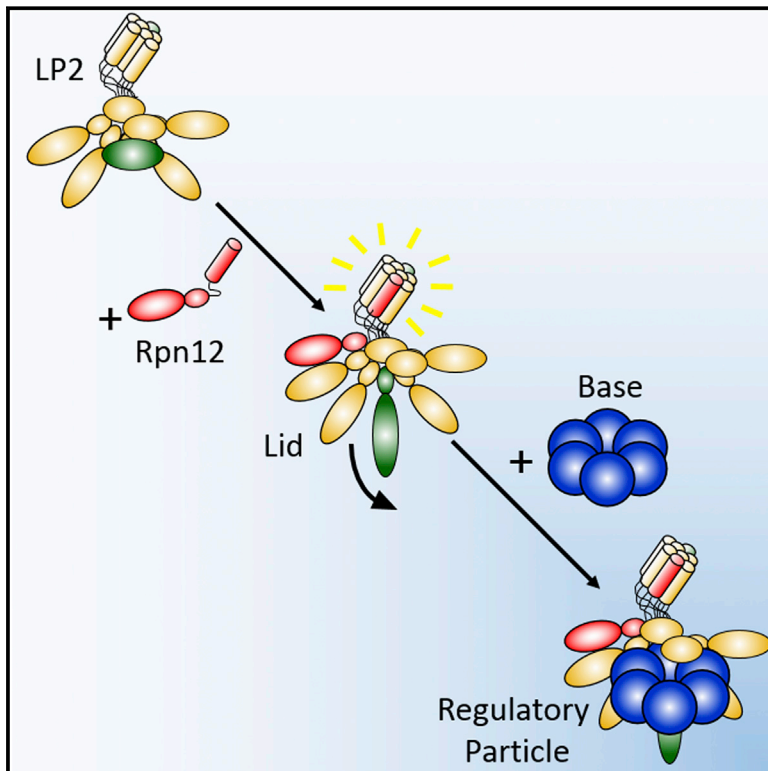
REFERENCES

- Abel, S., Chien, P., Wassmann, P., Schirmer, T., Kaever, V., Laub, M.T., Baker, T.A., and Jenal, U. (2011). Regulatory cohesion of cell cycle and cell differentiation through interlinked phosphorylation and second messenger networks. *Mol. Cell* 43, 550–560.
- Abel, S., Bucher, T., Nicollier, M., Hug, I., Kaever, V., Abel Zur Wiesch, P., and Jenal, U. (2013). Bi-modal distribution of the second messenger c-di-GMP controls cell fate and asymmetry during the *caulobacter* cell cycle. *PLoS Genet.* 9, e1003744.
- Ades, S.E. (2004). Proteolysis: Adaptor, adaptor, catch me a catch. *Curr. Biol.* 14, R924–R926.
- Battesti, A., and Gottesman, S. (2013a). Roles of adaptor proteins in regulation of bacterial proteolysis. *Curr. Opin. Microbiol.* 16, 140–147.
- Battesti, A., Hoskins, J.R., Tong, S., Milanesio, P., Mann, J.M., Kravats, A., Tsegaye, Y.M., Boudgour, A., Wickner, S., and Gottesman, S. (2013b). Anti-adaptors provide multiple modes for regulation of the RssB adaptor protein. *Genes Dev.* 27, 2722–2735.
- Bhat, N.H., Vass, R.H., Stoddard, P.R., Shin, D.K., and Chien, P. (2013). Identification of ClpP substrates in *Caulobacter crescentus* reveals a role for regulated proteolysis in bacterial development. *Mol. Microbiol.* 88, 1083–1092.
- Biondi, E.G., Skerker, J.M., Arif, M., Prasol, M.S., Perchuk, B.S., and Laub, M.T. (2006). A phosphorelay system controls stalk biogenesis during cell cycle progression in *Caulobacter crescentus*. *Mol. Microbiol.* 59, 386–401.
- Boudgour, A., Wickner, S., and Gottesman, S. (2006). Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in *Escherichia coli*. *Genes Dev.* 20, 884–897.
- Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., and Biondi, E.G. (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Syst. Biol.* 4, 52.
- Brötz-Oesterhelt, H., Beyer, D., Kroll, H.P., Endermann, R., Ladel, C., Schroeder, W., Hinzen, B., Raddatz, S., Paulsen, H., Henninger, K., et al. (2005). Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med.* 11, 1082–1087.
- Chien, P., Perchuk, B.S., Laub, M.T., Sauer, R.T., and Baker, T.A. (2007). Direct and adaptor-mediated substrate recognition by an essential AAA+ protease. *Proc. Natl. Acad. Sci. USA* 104, 6590–6595.
- Curtis, P.D., and Brun, Y.V. (2010). Getting in the loop: regulation of development in *Caulobacter crescentus*. *Microbiol. Mol. Biol. Rev.* 74, 13–41.
- Domian, I.J., Quon, K.C., and Shapiro, L. (1997). Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell* 90, 415–424.
- Dougan, D.A., Weber-Ban, E., and Bukau, B. (2003). Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Mol. Cell* 12, 373–380.
- Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B., and Jenal, U. (2009). Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes Dev.* 23, 93–104.

- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345.
- Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell Dev. Biol.* 19, 565–587.
- Guo, M.S., and Gross, C.A. (2014). Stress-induced remodeling of the bacterial proteome. *Curr. Biol.* 24, R424–R434.
- Iniesta, A.A., McGrath, P.T., Reisenauer, A., McAdams, H.H., and Shapiro, L. (2006). A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc. Natl. Acad. Sci. USA* 103, 10935–10940.
- Jenal, U., and Fuchs, T. (1998). An essential protease involved in bacterial cell-cycle control. *EMBO J.* 17, 5658–5669.
- King, R.W., Deshaies, R.J., Peters, J.M., and Kirschner, M.W. (1996). How proteolysis drives the cell cycle. *Science* 274, 1652–1659.
- Kirstein, J., Molière, N., Dougan, D.A., and Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nat. Rev. Microbiol.* 7, 589–599.
- Kobayashi, H., De Nisco, N.J., Chien, P., Simmons, L.A., and Walker, G.C. (2009). *Sinorhizobium meliloti* CpdR1 is critical for co-ordinating cell cycle progression and the symbiotic chronic infection. *Mol. Microbiol.* 73, 586–600.
- Kononova, A., Sogaard-Andersen, L., and Kroos, L. (2014). Regulated proteolysis in bacterial development. *FEMS Microbiol. Rev.* 38, 493–522.
- Lau, J., Hernandez-Alicea, L., Vass, R.H., and Chien, P. (2015). A phospho-signaling adaptor primes the AAA+ protease ClpXP to drive cell cycle regulated proteolysis. *Mol. Cell* 59, 104–116.
- Lori, C., Ozaki, S., Steiner, S., Böhm, R., Abel, S., Dubey, B.N., Schirmer, T., Hiller, S., and Jenal, U. (2015). Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. *Nature* 523, 236–239.
- Levchenko, I., Grant, R.A., Wah, D.A., Sauer, R.T., and Baker, T.A. (2003). Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. *Mol. Cell* 12, 365–372.
- Marczynski, G.T., and Shapiro, L. (2002). Control of chromosome replication in *caulobacter crescentus*. *Annu. Rev. Microbiol.* 56, 625–656.
- McGrath, P.T., Iniesta, A.A., Ryan, K.R., Shapiro, L., and McAdams, H.H. (2006). A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. *Cell* 124, 535–547.
- Ozaki, S., Schalch-Moser, A., Zumthor, L., Manfredi, P., Ebbensgaard, A., Schirmer, T., and Jenal, U. (2014). Activation and polar sequestration of PopA, a c-di-GMP effector protein involved in *Caulobacter crescentus* cell cycle control. *Mol. Microbiol.* 94, 580–594.
- Pini, F., De Nisco, N.J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., Mengoni, A., Bazzicalupo, M., Viollier, P.H., Walker, G.C., and Biondi, E.G. (2015). Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet.* 11, e1005232.
- Poindexter, J.S. (1981). The caulobacters: ubiquitous unusual bacteria. *Microbiol. Rev.* 45, 123–179.
- Quon, K.C., Marczynski, G.T., and Shapiro, L. (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* 84, 83–93.
- Radhakrishnan, S.K., Thanbichler, M., and Viollier, P.H. (2008). The dynamic interplay between a cell fate determinant and a lysozyme homolog drives the asymmetric division cycle of *Caulobacter crescentus*. *Genes Dev.* 22, 212–225.
- Radhakrishnan, S.K., Pritchard, S., and Viollier, P.H. (2010). Coupling prokaryotic cell fate and division control with a bifunctional and oscillating oxidoreductase homolog. *Dev. Cell* 18, 90–101.
- Rood, K.L., Clark, N.E., Stoddard, P.R., Garman, S.C., and Chien, P. (2012). Adaptor-dependent degradation of a cell-cycle regulator uses a unique substrate architecture. *Structure* 20, 1223–1232.
- Ryan, K.R., Judd, E.M., and Shapiro, L. (2002). The CtrA response regulator essential for *Caulobacter crescentus* cell-cycle progression requires a bipartite degradation signal for temporally controlled proteolysis. *J. Mol. Biol.* 324, 443–455.
- Schallies, K.B., Sadowski, C., Meng, J., Chien, P., and Gibson, K.E. (2015). *Sinorhizobium meliloti* CtrA stability is regulated in a CbrA-dependent manner and influenced by CpdR1. *J. Bacteriol.* 197, 2139–2149.
- Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., and Laub, M.T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol.* 3, 1770–1788.
- Smith, S.C., Joshi, K.K., Zik, J.J., Trinh, K., Kamajaya, A., Chien, P., and Ryan, K.R. (2014). Cell cycle-dependent adaptor complex for ClpXP-mediated proteolysis directly integrates phosphorylation and second messenger signals. *Proc. Natl. Acad. Sci. USA* 111, 14229–14234.
- Taylor, J.A., Wilbur, J.D., Smith, S.C., and Ryan, K.R. (2009). Mutations that alter RcdA surface residues decouple protein localization and CtrA proteolysis in *Caulobacter crescentus*. *J. Mol. Biol.* 394, 46–60.
- Thanbichler, M. (2010). Synchronization of chromosome dynamics and cell division in bacteria. *Cold Spring Harb. Perspect. Biol.* 2, a000331.
- Zhou, Y., Gottesman, S., Hoskins, J.R., Maurizi, M.R., and Wickner, S. (2001). The RssB response regulator directly targets sigma(S) for degradation by ClpXP. *Genes Dev.* 15, 627–637.

A Single α Helix Drives Extensive Remodeling of the Proteasome Lid and Completion of Regulatory Particle Assembly

Graphical Abstract



Authors

Robert J. Tomko, Jr., David W. Taylor, Zhuo A. Chen, Hong-Wei Wang, Juri Rappsilber, Mark Hochstrasser

Correspondence

robert.tomko@med.fsu.edu (R.J.T.), mark.hochstrasser@yale.edu (M.H.)

In Brief

A single alpha helix from the final subunit that incorporates into the proteasomal lid triggers a large-scale conformational switch that enables subsequent assembly of the lid and base, suggesting a general paradigm for hierarchical assembly of macromolecular complexes similar to that of virus particles.

Highlights

- First in vitro reconstitution of RP assembly with completely recombinant components
- Electron microscopy and cross-linking reveal massive remodeling of a lid precursor
- Remodeling of the lid relieves steric clash with the RP base to promote RP assembly
- Lid remodeling can be triggered by a single C-terminal α helix in the Rpn12 subunit



A Single α Helix Drives Extensive Remodeling of the Proteasome Lid and Completion of Regulatory Particle Assembly

Robert J. Tomko, Jr.,^{1,7,*} David W. Taylor,² Zhuo A. Chen,³ Hong-Wei Wang,⁴ Juri Rappsilber,^{3,5,6} and Mark Hochstrasser^{1,6,*}

¹Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, USA

²California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, CA 94720-3200, USA

³Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, King's Buildings, Max Born Crescent, Mayfield Road, Edinburgh EH9 3BF, Scotland

⁴Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, PRC

⁵Department of Bioanalytics, Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany

⁶Co-senior author

⁷Present address: Department of Biomedical Sciences, Florida State University College of Medicine, Tallahassee, FL 32306-4300, USA

*Correspondence: robert.tomko@med.fsu.edu (R.J.T.), mark.hochstrasser@yale.edu (M.H.)

<http://dx.doi.org/10.1016/j.cell.2015.09.022>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

SUMMARY

Most short-lived eukaryotic proteins are degraded by the proteasome. A proteolytic core particle (CP) capped by regulatory particles (RPs) constitutes the 26S proteasome complex. RP biogenesis culminates with the joining of two large subcomplexes, the lid and base. In yeast and mammals, the lid appears to assemble completely before attaching to the base, but how this hierarchical assembly is enforced has remained unclear. Using biochemical reconstitutions, quantitative cross-linking/mass spectrometry, and electron microscopy, we resolve the mechanistic basis for the linkage between lid biogenesis and lid-base joining. Assimilation of the final lid subunit, Rpn12, triggers a large-scale conformational remodeling of the nascent lid that drives RP assembly, in part by relieving steric clash with the base. Surprisingly, this remodeling is triggered by a single Rpn12 α helix. Such assembly-coupled conformational switching is reminiscent of viral particle maturation and may represent a commonly used mechanism to enforce hierarchical assembly in multisubunit complexes.

INTRODUCTION

The eukaryotic proteome is constantly remodeled to control cell function. This remodeling depends heavily on protein degradation by the ubiquitin-proteasome system (UPS) (Finley, 2009; Ravid and Hochstrasser, 2008). The UPS consists of a cascade of enzymes that catalyze the transfer of the protein ubiquitin to protein substrates destined for degradation by the 26S proteasome. The proteasome is a 2.5 MDa ATP-dependent protease complex composed of a barrel-shaped proteolytic core particle

capped on its open ends by 19S regulatory particles (Tomko and Hochstrasser, 2013). Under certain conditions, the regulatory particles (RP) can be divided into subcomplexes called the lid and base (Glickman et al., 1998). The lid consists of nine subunits, Rpn3, 5–9, 11, 12, and Sem1. The base contains a ring of six AAA+ family ATPases, Rpt1–6, and the non-ATPase subunits Rpn1, 2, 10, and 13. The lid removes the ubiquitin tag from substrates, whereas the base uses mechanical energy derived from ATP hydrolysis to unfold substrates and insert them into the core particle (CP) for destruction (Nyquist and Martin, 2014; Tomko and Hochstrasser, 2013).

The proteasome consists of at least 33 distinct subunits. Its assembly is a tightly coordinated process that relies both on dedicated extrinsic assembly chaperones and intrinsic features of the subunits themselves. Whereas both the CP and RP base depend heavily on assembly chaperones for efficient assembly, the lid can assemble independently of any additional eukaryotic factors (Fukunaga et al., 2010; Tomko and Hochstrasser, 2011, 2014). Lid biogenesis appears to follow a defined assembly sequence that culminates with addition of the Rpn12 subunit to a nearly complete lid intermediate consisting of Rpn3, 5–9, 11, and Sem1, called lid particle 2 (LP2) (Tomko and Hochstrasser, 2011). Recombinant lid forms efficiently in the absence of the base or CP (Lander et al., 2012; Tomko and Hochstrasser, 2014). Importantly, LP2 is unable to participate in further 26S proteasome assembly unless Rpn12 is present to complete lid formation (Tomko and Hochstrasser, 2011).

Incorporation of Rpn12 into LP2 licenses the resultant complex for assembly into full proteasomes, but the molecular mechanism underlying this critical function for Rpn12 has remained obscure (Tomko and Hochstrasser, 2011). In the mature 26S proteasome, Rpn12 occupies a peripheral position within the RP (Figures 1A and 1B) and contributes minimally to the interface between lid and base (Matyskiela et al., 2013; Unverdorben et al., 2014). Despite this, the LP2 intermediate has no detectable affinity for the base in the absence of the latter subunit (Tomko and Hochstrasser, 2011).

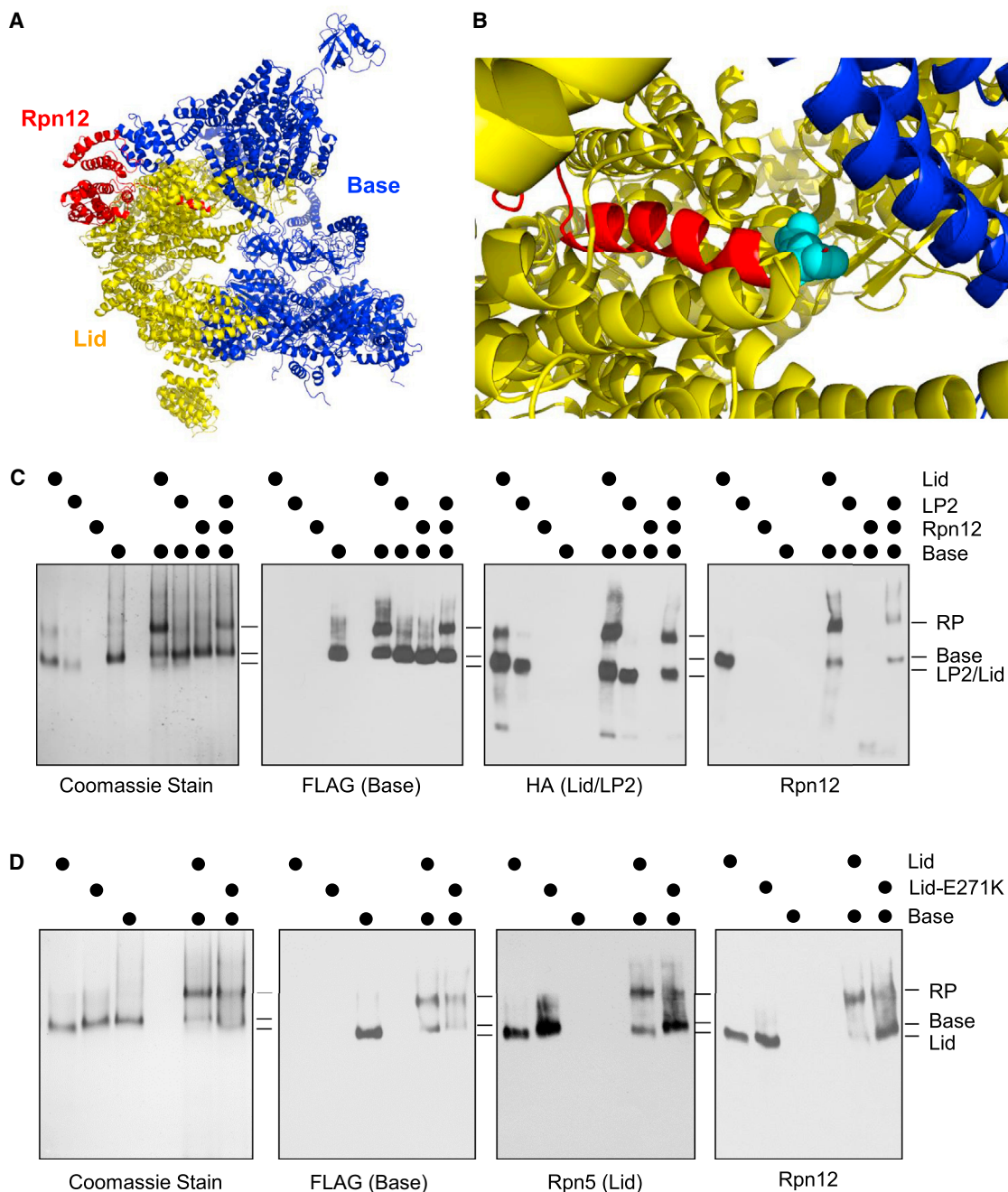


Figure 1. Reconstitution of the 19S Regulatory Particle Using Purified Components

(A) Pseudoatomic model of the proteasome regulatory particle (RP) (from PDB 4CR2) indicating the position of Rpn12 (red) with respect to (yellow) and base (blue) subunits. For clarity, the CP, as well as the RP subunits Rpn1 and Rpn10, is omitted.

(B) The Rpn12-Glu271 residue (shown in cyan) lies near the lid-base interface. CP, Rpn1, and Rpn10 have been omitted as in (A).

(C) RP assembly depends on Rpn12. The indicated recombinant components (1 μ M) were incubated together in the presence of 1 mM ATP and 10 μ M recombinant Rpn10 for 20 min at 30°C before separation by native PAGE. Gels were either stained with Coomassie brilliant blue or transferred to PVDF membranes followed by immunoblotting with antibodies to FLAG (on the Rpt1 base subunit), HA (on the Rpn7 LP2 subunit), or Rpn12.

(D) The rpn12-E271K mutation weakens lid-base interaction without interfering with lid formation. The indicated components (1 μ M) were incubated with 1 mM ATP and 10 μ M Rpn10 before analysis as in (C).

See also Figure S1.

Proteasomes have recently emerged as important targets for the treatment of certain cancers (Crawford and Irvine, 2013). All of the anti-proteasome drugs currently FDA approved or in clinical trials are inhibitors of the CP active sites (Dou and Zonder, 2014), but drug resistance is already a significant problem (Cheriyath et al., 2007; Richardson et al., 2003). Interference with proteasome assembly could provide a valuable alternative strategy for chemotherapy. As incorporation of Rpn12 is obligatory for subsequent RP assembly, interference with this final step of lid assembly would be an attractive target for such therapies. However, this will require more detailed knowledge of lid biogenesis and a better understanding of how Rpn12 regulates RP assembly. Thus, a central challenge in studying proteasome biogenesis, as well as that of many other multisubunit complexes, is to understand precisely how hierarchical assembly events are enforced *in vivo*.

Single-particle electron microscopy (EM) and mass spectrometry (MS) are powerful tools to investigate large macromolecular structures. Although these approaches have been used extensively with the fully assembled 26S proteasome (Lander et al., 2012; Sharon et al., 2006, 2007; Unverdorben et al., 2014; Witt et al., 2006), little information is available on the structures of proteasome assembly intermediates (Kock et al., 2015). Application of structural methods to assembly intermediates could provide substantial insight into the molecular mechanisms guiding proteasome biogenesis.

Quantitative cross-linking/MS (QCLMS) has recently emerged as a means to provide insights into dynamic aspects of protein structures (Fischer et al., 2013). An early application revealed structural changes within a multi-protein complex upon phosphorylation (Schmidt et al., 2013). Here, using a newly established workflow, we apply QCLMS to both the LP2 assembly intermediate and full lid; this has revealed multiple conformational changes, including the repositioning of the Rpn8-Rpn11 deubiquitinating enzyme (DUB) module. In addition, by comparing EM-derived molecular models of LP2, the free lid and the lid docked within the 26S proteasome, we show that incorporation of Rpn12 results in large-scale conformational remodeling of the nascent lid from a compact “closed” state in LP2, which is autoinhibited for base binding, to an open, more flexible state that can pair with the base. Remarkably, the highly conserved C-terminal α helix of Rpn12 is sufficient to promote conversion of LP2 to a lid-like conformation, enabling assembly of LP2 into the RP. These findings rationalize previous biochemical and genetic data indicating a critical role for Rpn12 in RP assembly, and provide novel insights into chaperone-independent proteasome assembly. Coupling small, local structural switches to large conformational changes is reminiscent of certain steps in virus assembly. Such locally induced conformational rearrangements are likely to be used in the hierarchical assembly of other multisubunit complexes, such as snRNPs, group II chaperonins, and ribosomes.

RESULTS

Reconstitution of Proteasome RP Assembly from Recombinant Constituents

Previously, we found that purified yeast LP2 assembled into full proteasomes when added to a yeast extract containing

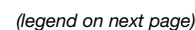
wild-type Rpn12, but not one with an Rpn12 truncation lacking C-terminal residues 212–274 (Figure S1) (Tomko and Hochstrasser, 2011). This could be due to a requirement for full-length Rpn12 or from an inhibitory effect of the truncated Rpn12 protein. Assembly might also require other factors in the extracts such as molecular chaperones. Thus, we sought a system of reduced complexity to analyze RP assembly. Toward this end, we established an *in vitro* RP assembly assay that consisted exclusively of bacterially expressed recombinant proteins.

Interaction of recombinant LP2 and a recombinant base precursor containing the Rpn14, Nas6, and Hsm3 assembly chaperones (hereafter called base) (Beckwith et al., 2013) was entirely dependent on Rpn12 (Figure 1C). Notably, Rpn12 could be supplied as part of the lid or added ectopically to the LP2 and base mixture with comparable results. In the pseudoatomic structure of the 26S proteasome determined by cryo-EM (PDB 4CR2), Rpn12 Glu271 lies near the interface between the lid and base (Figure 1B). As previously observed in yeast cells (Tomko and Hochstrasser, 2011), mutation of this highly conserved residue to lysine also impairs lid-base attachment in the *in vitro* system, but it does not compromise the stability of Rpn12 within the lid (Figure 1D). These findings support our previous observations that Rpn12 incorporation controls lid-base association in yeast and is regulated at least in part by the C-terminal helix of Rpn12.

QCLMS Reveals Local LP2-Lid Structural Changes

Rpn12 occupies a peripheral position within the RP and provides a very small fraction of the total interface between the lid and base complexes. This suggests that Rpn12 may instead promote lid-base joining primarily by inducing conformational changes in the assembling lid. To help determine how Rpn12 mediates lid-base joining, we first utilized QCLMS (Fischer et al., 2013) to identify protein cross-links that were unique or enriched in LP2 or lid, with the expectation that these differences reflect differences between the solution state conformations of these two particles. We purified each of the complexes from yeast (Figures S2A and S2B) and subjected them to cross-linking with non-deuterated or deuterated forms of the amine-reactive cross-linker BS³, followed by protease digestion, 1:1 sample mixing, and LC-MS/MS analysis (Figure 2A). We used conditions identified in preliminary experiments that yielded only intra-complex cross-links (Figures S2C and S2D). The cross-linking experiments were done twice but with the non-deuterated and deuterated BS³ swapped between LP2 and lid samples; this replicate analysis with label-swap controlled for any labeling bias and ensured accurate quantitation, especially for cross-links that are unique in either lid or LP2 (Figures S2E–S2G).

We identified and quantified 122 unique cross-links (Table S1). Among them, 81 were quantified consistently in label-swap replicates. Most of these 81 cross-links showed nearly equal prevalence in the lid and LP2 samples (i.e., were centered on $\log_2 = 0$; Figure 2B, i, and Figures S2E and S2F), which is consistent with 1:1 sample mixing (Figure 2B). However, one cross-link was unique to lid, and three were unique to LP2 (Figure 2B, ii). In addition, 16 of the 77 cross-links observed for both lid and LP2 were significantly enriched in either lid or LP2 (“Significance A” test, Perseus version 1.4.1.2 [Cox and Mann, 2008]; $p < 0.05$). Of these 16 cross-links, 15 changed more than 4-fold in at least



one of the replicated analyses (Figure S2F). Importantly, none of these cross-links directly involved Rpn12 or were in positions expected to be sterically occluded by Rpn12 based on the pseudoatomic structure of the 26S proteasome, strongly suggesting that these differences resulted from structural variance between the two particles.

Examination of the cross-links enriched in either LP2 or the lid suggested two major conformational differences between the two particles. The first was the relative positioning of the Rpn8-Rpn11 module (regions A1 [Rpn8 residues 299–309], A2 [Rpn11 residue 267], and B [Rpn7 residue 2], Figures 2C and 2D) within the complexes, and the second was the positioning of the Sem1 N terminus, the N-terminal α -helical domain of Rpn3, and the C-terminal α -helix of Rpn7 relative to one another (regions C1 [Sem1 residues 2–26], C2 [Rpn7 residues 416–426], and C3 [Rpn3 residues 299–389], Figure 2C). The latter three subunits yielded a web of cross-links that was selectively enriched in LP2, suggesting that these regions were more closely packed in LP2 (pink lines between regions C1, 2, and 3, Figure 2C). The enrichment of cross-links between the C-terminal helix of Rpn7 and residues in the N-terminal extension of Rpn3 (Figure S1) suggests that the helical bundle, or at least the C-helix of Rpn7, may be compacted against the body of Rpn3 and/or Rpn7 in LP2, and that it is repositioned upon incorporation of Rpn12. As Rpn12 forms an extensive interface with Rpn3 and appears to make little to no contact with Rpn7 in the 26S proteasome structure (Matyskiela et al., 2013; Unverdorben et al., 2014), it is likely that these changes are mediated through Rpn3.

Regarding the Rpn8-Rpn11 module, cross-links between Rpn8 and the C-terminal helix of Rpn9 (Figure S1) were enriched or unique to LP2 (Figure 2C), and an LP2-specific cross-link between Rpn8 and the extreme N terminus of Rpn7 was observed (red line between A1 and B, Figure 2C). In contrast, we observed lid-enriched cross-linking between the C terminus of Rpn8 and the C-terminal helix of Rpn7, the N-terminal region of Rpn3, and the N-terminal region of Sem1 (regions C1 and C3, light blue lines, Figure 2C). Further, a cross-link between the C-terminal helix of Rpn11 and the N-terminal region of Sem1 was detected only in the lid (regions A2 and C1, dark blue line, Figure 2C). A series of common cross-links between Rpn8 and Rpn11 suggests that the module itself is unlikely to vary substantially in structure between LP2 and lid.

Taken together, the enrichment of cross-linking in different regions of the lid and the loss or depletion of cross-linking to Rpn9 suggest that the Rpn8-Rpn11 module undergoes substantial movement upon incorporation of Rpn12 (Figure 2D). We suggest that a rigid-body rotation of the Rpn8-Rpn11 dimer (modeled in Figure 2E) may occur upon incorporation of Rpn12 that would account for the observed changes in cross-linking.

The N Terminus of Rpn5 Is Highly Flexible in the Undocked Lid

Currently, the only structural information on the isolated lid is from negative-stain EM (Lander et al., 2012). However, this structure is not at sufficiently high resolution to aid in visualizing our cross-linking data. To provide insight into the solution conformation of the free lid, we mapped our cross-link network for the isolated lid onto the lid within the cryo-EM structure of the 26S proteasome (Unverdorben et al., 2014). 61 cross-links could be mapped onto this structure (Figure 3, Data S1, and Table S1). Surprisingly, 26 of the 61 cross-links spanned distances greater than the calculated maximum distance between α -carbons in BS³-cross-linked residues (based on spacer length of BS³ and side-chain lengths of cross-linked residues; Supplemental Experimental Procedures) (Figure 3B), suggesting that the free lid is much more flexible than the docked lid. Two clusters of over-length inter-subunit cross-links were apparent. The first consisted of five cross-links between Rpn9 and either Rpn8 or Rpn11, indicating that the Rpn8/Rpn11 dimer is positioned closer to Rpn9 in the free lid than in the mature proteasome. Importantly, together with the results of our LP2-lid comparison, this further supports a model in which Rpn8/Rpn11 undergoes substantial movements during RP assembly.

The second cluster involved 19 cross-links between the N-terminal extension of Rpn5 and either itself or other lid subunits, with several of these cross-linked residues being over 100 Å apart in the 26S proteasome. Further, cross-links involving Rpn5-K108 were observed with residues from four subunits (Rpn3, Rpn6, Rpn8, and Rpn11) that are, in some cases, very far from one another in the proteasome-docked lid, suggesting substantial mobility of this domain of Rpn5 in the free lid. This was further supported by the observation that over-length intra-subunit cross-links were observed only within Rpn5 and not within any other subunit. Taken together, these data strongly suggest that the lid Rpn5 N-terminal domain is highly flexible

Figure 2. Quantitative Cross-linking/Mass Spectrometry Reveals Local Differences in LP2 and Lid Structures

- (A) Overview of quantitative cross-linking/MS analysis of LP2 and lid complexes.
- (B) *i*: fold changes [$\log_2(\text{LP2}/\text{lid})$] of 81 quantified cross-links (consistently quantified in both forward- and reverse-labeled experiments) were plotted against their signal intensities. These quantified cross-links were divided into five subgroups and colored accordingly—three are unique to the LP2 sample (red), one is unique to the lid sample (blue); among 77 cross-links observed in both samples, 61 common cross-links showed no significant difference in yields in the LP2 and the lid samples (gray), while seven cross-links are significantly upregulated in the LP2 sample ($p < 0.05$, magenta) and nine in the lid sample ($p < 0.05$, cyan). *ii*: observed MS signals of supporting cross-linked peptides of “lid unique” cross-link 267^{RPN11}-20^{SEM1} (left), “common” cross-link 404^{RPN5}-150^{RPN11} (middle), and “LP2 unique” cross-link 2^{RPN7}-300^{RPN8} (right).
- (C) Cross-linking network of lid/LP2 subunits. Proteins are shown as bars, and cross-links are shown as lines that are colored according to their quantified subgroups (described in B). LP2-unique, LP2-enriched, lid-unique, and lid-enriched cross-links, which suggest conformational differences between LP2 and lid, mainly involve three regions marked as A, B and C.
- (D) Relative cross-link enrichments suggest a repositioning of the Rpn8-Rpn11 heterodimer within the lid structure (derived from PDB 4CR2) upon Rpn12 incorporation that may be important for subsequent RP assembly steps. Regions A, B, and C are highlighted accordingly. Regions C1 and B that are not present in 4CR2 are located based on their cross-linking partners in the model.
- (E) A rigid-body rotation of the Rpn8-Rpn11 upon Rpn12 incorporation may account for the observed change in the Rpn8-Rpn11 cross-link network.

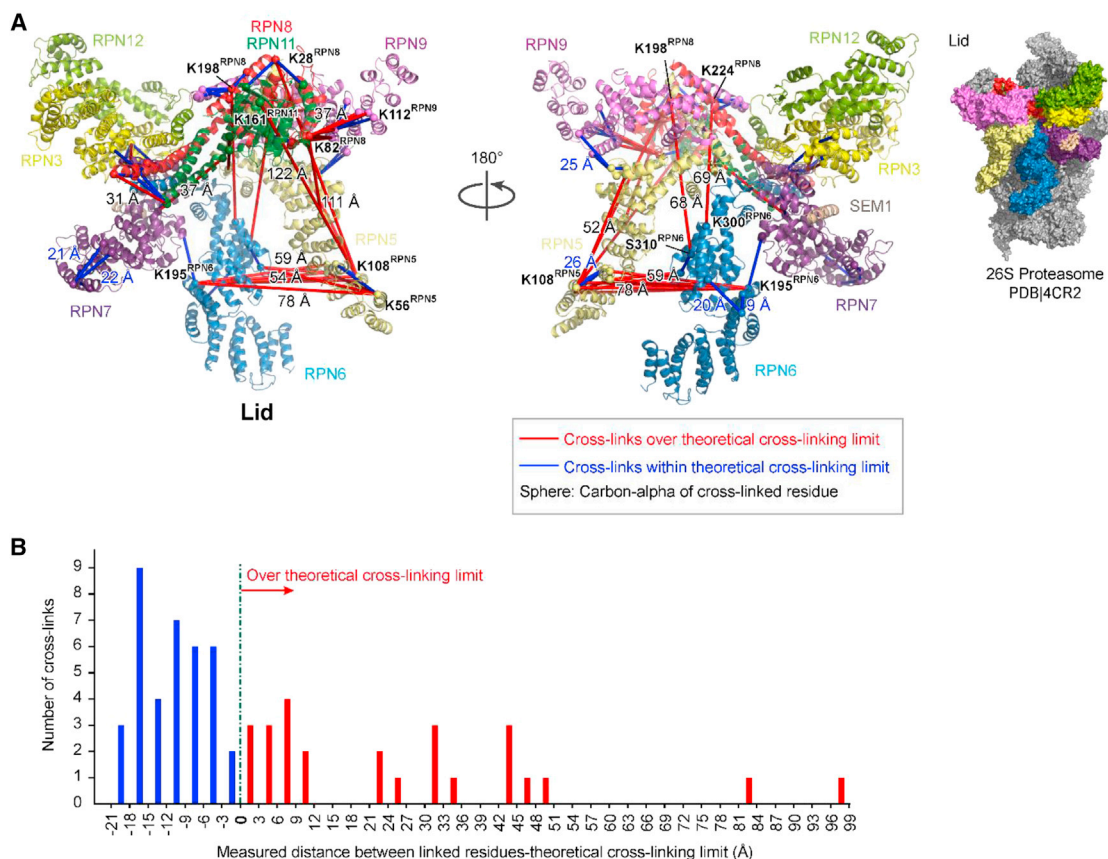


Figure 3. Mapping Lid Cross-links onto the Lid in the 26S Pseudoatomic Model Reveals Subunit Flexibility in the Undocked Lid

(A) Lid cross-links (61) detected in our QCLMS analysis are depicted on the 26S proteasome pseudoatomic structure (PDB 4CR2). Cross-links are displayed as straight lines connecting cross-linked residues, and distances between them are measured between C α atoms. Cross-links are displayed only when both cross-linked residues are present in the proteasome pseudoatomic structure. Cross-links exceeding the predicted limit are shown in red, whereas those within the limit are shown in blue.

(B) The differences in length between the 61 cross-links mapped onto PDB 4CR2 in (A) and the theoretical cross-linking limit (e.g., 27 Å for lysine-lysine cross-links, [Supplemental Experimental Procedures](#)) are shown. Distances are between the C α atoms of cross-linked residues.

See also [Figure S2](#).

and free to undergo substantial movement in solution ([Lander et al., 2012](#)). Incorporation of the lid into the 26S proteasome likely stabilizes Rpn5 in the extended conformation via interaction of its N terminus with the base and CP and is accompanied by a repositioning of the Rpn8/Rpn11 module away from Rpn9.

EM Structure of LP2 Reveals a Compact Intermediate

We next investigated the architecture of LP2 by negative-stain EM and single-particle analysis for comparison to the negative stain EM structure of the isolated lid (EMD-1993) ([Lander et al., 2012](#)). Raw micrographs of LP2 showed monodisperse particles with dimensions ≈ 200 Å \times 250 Å ([Figure S3A](#)). Reference-free two-dimensional (2D) alignment and classification yielded class averages with strong similarities to the previously determined structure of the isolated lid ([Figure 4A](#)). As expected, LP2 lacked density for Rpn12, but surprisingly, also lacked the prominent N-terminal extension of Rpn6 that was previously observed in the isolated lid ([Lander et al., 2012](#)). We saw this absence of density for the N-terminal portion of Rpn6 both for LP2 purified

from yeast, as well as recombinant LP2 produced in *E. coli* ([Figures S3G and S3H](#)).

To address whether the observed LP2-lid differences were a result of sample preparation differences, we repeated the EM analysis with the full lid. 2D class averages of lid purified from yeast showed a density for the N-terminal extension of Rpn6 ([Figure 4A](#)). Additionally, a three-dimensional (3D) reconstruction of lid obtained with our complexes was nearly identical to that determined by [Lander et al. \(2012\)](#) ([Figures 4B, S3C, and S3D](#)), indicating that sample preparation was not responsible for the structural differences between LP2 and lid. Using iterative back-projection refinement of LP2 particles with the structure of lid low-pass filtered to 60 Å as a starting model, we obtained a 3D reconstruction of LP2 at 16 Å resolution, showing the clear domain architecture of the complex ([Figures 4B and S3E–S3G](#)).

The quality of our LP2 structure allowed facile automated segmentation based on visual inspection and comparison to the previously determined lid density ([Figure 4B](#)). In agreement with our 2D analysis and protein cross-linking, the overall subunit

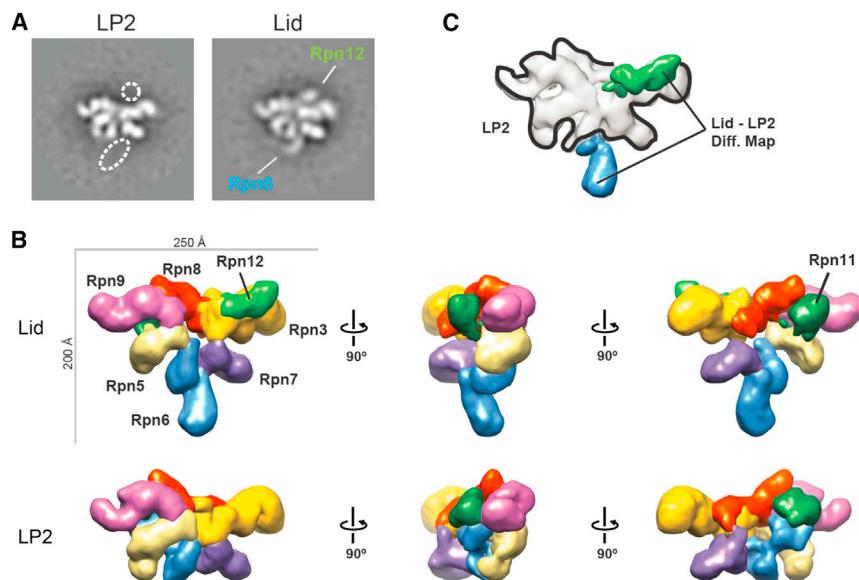


Figure 4. Electron Microscopy of LP2 and Lid Purified from Yeast Reveals Closed and Open Configurations, Respectively

(A) The N-terminal domain of Rpn6 adopts an extended conformation in lid, but not in LP2. Representative 2D class averages of LP2 or lid indicating the positions of Rpn12 and the Rpn6 N terminus are shown. The width of the boxes is ~ 430 Å.

(B) 3D reconstructions of the lid (EMD-1993) and LP2 (this study) are shown.

(C) Difference-density map of lid and LP2 indicates that the two major regions of variant density are those for Rpn12 and the N-terminal domain of Rpn6.

See also Figure S3.

architectures of LP2 and lid were very similar. In LP2, the Rpn6 N-terminal helical domain (Rpn6-N) is folded into a cavity formed by Rpn3, Rpn7, Rpn8, and Rpn11. Using 3D difference mapping between our LP2 structure and the lid structure from Lander et al. (2012), we unambiguously assigned the most prominent changes in density in the lid to Rpn12 and Rpn6-N (Figure 4C). This observation prompted us to reexamine our QCLMS data. Although the Rpn6-N conformation seen by EM in LP2 was not overtly reflected in LP2-specific cross-linking of this domain (Figure 3), an LP2-unique cross-link was in fact observed between Rpn6-Lys195 and Rpn11-Lys12 (Table S1). This particular amino acid of Rpn11 is not modeled in the cryo-EM structure (PDB: 4CR2), but extrapolation from the most proximal modeled Rpn11 residue (Thr23) indicates that the cross-link between Rpn6 and Rpn11 must be well over the maximal predicted cutoff. In further support of the compact conformation of LP2 observed by EM, two overlength cross-links were observed between Rpn6-N and Rpn8 (Rpn6-Lys300 to Rpn8-Thr224 and Rpn6-Ser310 to Rpn8-Lys198) for both the lid and LP2 (Figure 3A and Table S1). Formation of these cross-links in the lid, as well as LP2, suggests that the lid may dynamically sample more compact conformations.

LP2 Is Sterically Incompatible with Positioning in the Mature 26S Proteasome

In the 26S proteasome, the Rpn6 N-terminal helical domain is extended and makes contact with subunits of the RP base and CP (Lander et al., 2012; Pathare et al., 2012). Since this domain is folded into the core of LP2, it raised the possibility that adoption of this conformation produced a steric conflict with the base that would prevent incorporation of LP2 into the assembling RP prior to Rpn12 assimilation. By docking the previously determined EM structure of the isolated lid onto the lid subunits of the pseudoatomic model of the 26S and then overlaying our LP2 structure onto this structure, we were able to examine whether LP2 could incorporate into the protea-

some without steric conflict (Figure 5A). In the EM structure of the 26S proteasome, Rpn3 and Rpn9 curve inward toward the center of the RP and make direct contacts with Rpn2 and Rpn10, respectively. In

the LP2 model, Rpn3 and Rpn9 are not curved inward, which could contribute to the apparent poor affinity of LP2 for the base. Further, in LP2, Rpn6-N cannot make the contacts with the core particle and base ATPase ring that are observed for this domain in the mature 26S proteasome. Rather, the model displayed clear steric clashes between the LP2 Rpn6-N conformation and segments within the base ATPase subunits Rpt3, Rpt4, and Rpt6 (Figure 5A, inset). This strongly implies that the closed conformation observed for LP2 is an autoinhibited conformation that prevents lid-base association until Rpn12 incorporation. This model would rationalize the behavior of LP2 in vitro (Figure 1) and in yeast (Tomko and Hochstrasser, 2011).

Addition of Purified Rpn12 to LP2 Is Sufficient to Drive Lid Maturation

If Rpn12 relieves an autoinhibited conformation of LP2 to promote lid-base joining, then adding purified Rpn12 to LP2 should promote structural changes in LP2 that result in a lid-like structure. We treated purified recombinant LP2 with an equimolar concentration of recombinant Rpn12 and examined the structure of the resultant complex by EM. Consistent with a role for Rpn12 in maturation of the assembling lid, 2D class averages indicated that Rpn12 addition caused extension of Rpn6-N to a conformation similar to that observed in the purified full lid (Figure 5B). Thus, incorporation of Rpn12 drives a large-scale structural rearrangement of the assembling lid that relieves the autoinhibitory conformation of the Rpn6 N terminus.

Using two complementary approaches, we next addressed whether relieving the steric conflict between the N-terminal domain of Rpn6 and the base was sufficient to drive RP assembly in the absence of Rpn12. In the first, we appended the bulky GFP protein to the N terminus of Rpn6, which should physically clash with surrounding subunits in the closed conformation and potentially induce the extended conformation for Rpn6-N. In the

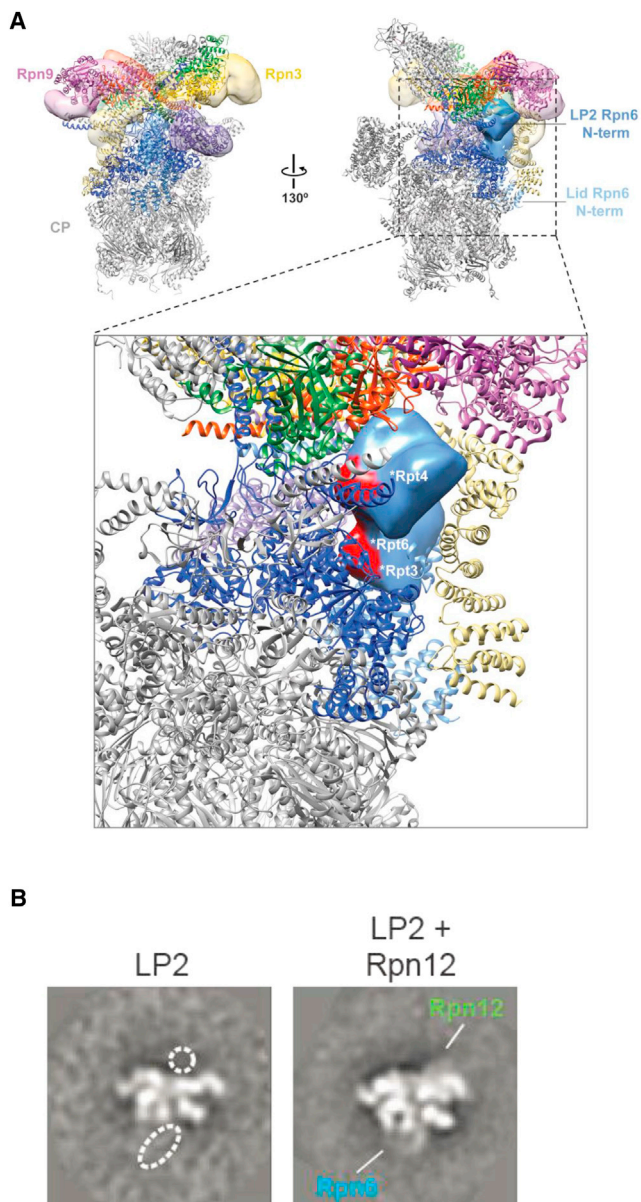


Figure 5. Rpn12 Addition Relieves an Autoinhibitory Conformation of the Rpn6 N Terminus

(A) The LP2 3D reconstruction (transparent surface) was modeled onto the pseudoatomic model of the 26S proteasome (PDB 4CR2) using the winged-helix horseshoe as a guide point, revealing steric clash (red surface) between the N-terminal domain of Rpn6 in LP2 (medium blue surface) and Rpt3, 4, and 6 (dark blue ribbons) of the base (inset). The corresponding position of the N-terminal domain of Rpn6 in the lid is also indicated (light blue ribbons). Rpn3 (yellow) and Rpn9 (magenta) are highlighted to show differences between LP2 and the docked lid.

(B) Addition of recombinant Rpn12 to purified recombinant LP2 triggers extension of the Rpn6 N terminus. The class averages are labeled as in Figure 4A. The width of the boxes is ~430 Å.

second approach, we engineered N-terminally truncated forms of Rpn6 that were anticipated to lack the putative autoinhibitory domain (Figure 6A). Both the GFP-Rpn6 fusion and N-terminal

Rpn6 truncation proteins assembled efficiently into LP2 when coexpressed with the other LP2 subunits in *E. coli* (Figures S4A and S4B), suggesting that Rpn6-N is dispensable for LP2 assembly. As predicted, appending GFP to the N terminus of Rpn6 resulted in a conformation of Rpn6-N within LP2 that appeared similar to its extended conformation in the full lid (Figures 6B and S4C). Importantly, each of the mutant forms of LP2 assembled into the RP when recombinant Rpn12 was provided (Figure S4D), indicating that they were functional for RP assembly. In contrast, neither GFP-LP2 (Figure 6C) nor LP2 harboring truncations of Rpn6 (Figure 6D) assembled into an RP-like structure in the absence of Rpn12. Therefore, repositioning or removal of Rpn6-N is insufficient to promote LP2-base joining. This indicates that Rpn12 binding to LP2 triggers additional changes in the lid precursor that are necessary for RP formation.

Rpn12 C-Terminal Helix Engagement of the Lid Helical Bundle Licenses RP Assembly

The C terminus of Rpn12 forms an α helix that inserts into a pronounced groove within a helical bundle formed by C-terminal helices from all the other large lid subunits (Estrin et al., 2013). This helix of Rpn12 is important both for the completion of lid assembly and engagement of the lid with the base (Tomko and Hochstrasser, 2011). The complex interface between the helical bundle and the Rpn12 helix would in principle allow Rpn12 to sense and transmit signals through multiple lid subunits. We therefore asked whether provision of the Rpn12 C-terminal helix alone was sufficient to drive lid-base association.

A fluorescently labeled peptide composed of Rpn12 residues 254–272 (N12pep) bound to purified LP2 in vitro and could be competed off with excess recombinant Rpn12, indicating that the peptide bound LP2 in a manner similar to the full-length protein (Figure 7A). Indeed, addition of N12pep to purified LP2 triggered conformational extension of the Rpn6 N-terminal domain (Figure 7B), as had been observed with full-length recombinant Rpn12 (Figure 5B). Most importantly, N12pep promoted LP2-base joining when LP2 and N12pep were supplied in molar excess to the base, as revealed by a migration shift of the base on native gel blots (Figure 7C). Assembly into the RP was not as efficient as it was with recombinant Rpn12, suggesting that other parts of Rpn12 promote assembly, possibly by stabilizing the Rpn12 C-terminal helix. RP formation was specific to N12pep, as identical concentrations of two unrelated peptides had no effect on LP2-base joining (Figure 7D). In summary, our data indicate that engagement of the nascent lid helical bundle by the Rpn12 C-terminal helix is the key event triggering conformational remodeling of the assembling lid and subsequent lid-base joining.

DISCUSSION

Hierarchical assembly has emerged as a theme in the regulation of proteasome biogenesis (Tomko and Hochstrasser, 2013). In yeast and mammals, most evidence suggests that the proteasome lid assembles completely before making stable contact with the base complex or its subunits (Fukunaga et al., 2010; Thompson et al., 2009; Tomko and Hochstrasser,

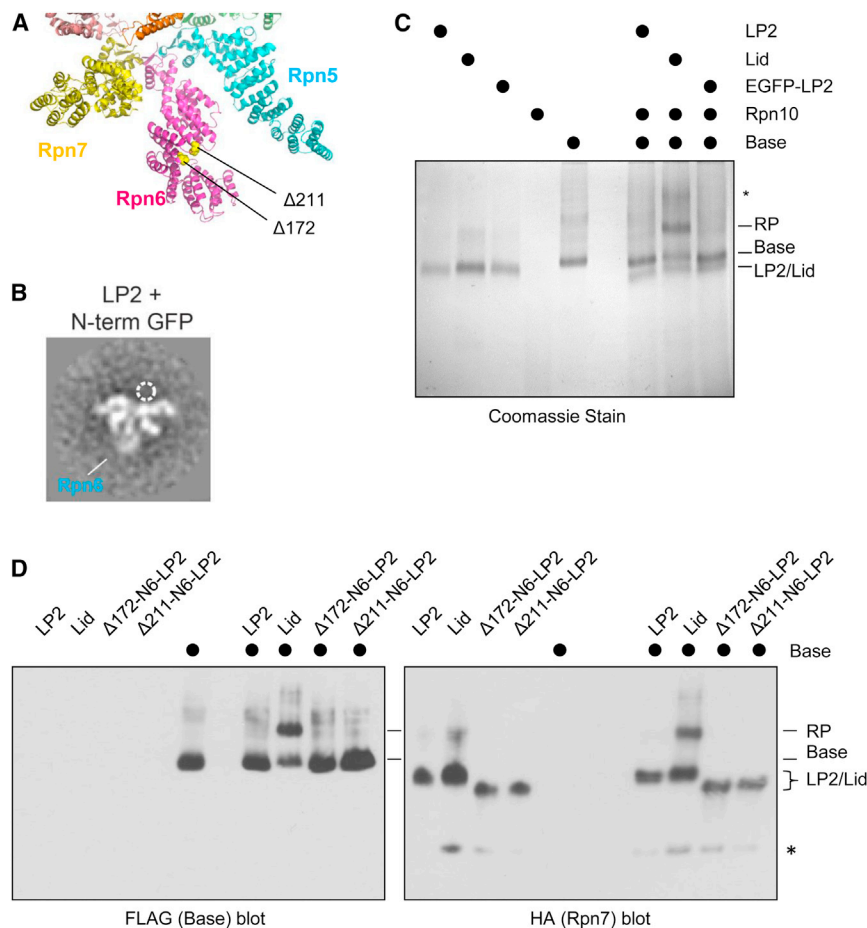


Figure 6. Relief of Steric Conflict Imposed by the Rpn6 N Terminus Is Not Sufficient to Drive Binding of LP2 to the Base

(A) The positions of truncations to Rpn6 are shown on the pseudoatomic structure of the lid from PDB 4CR2. The base and core particle subunits are omitted for clarity. The positions of Rpn6 amino acid 173 and Rpn6 amino acid 212 (the first endogenous residues present in the $\Delta 172$ and $\Delta 211$ -rpn6 truncation mutants, respectively) are shown as yellow spheres.

(B) Fusion of GFP to the N terminus of Rpn6 within LP2 causes the N terminus to assume a conformation similar to that observed for the lid. The width of the box is ~ 430 Å.

(C) Addition of LP2 containing a GFP-Rpn6 fusion to the base precursor is not sufficient to trigger LP2-base association. Untagged LP2, LP2 with N-terminally GFP-tagged Rpn6 (GFP-LP2), or lid (1 μ M each) were added to 1 μ M base and 10 μ M Rpn10 before analysis by native PAGE. Asterisk, poorly resolved or potentially aggregated protein(s). (D) Addition of LP2 containing truncated forms of Rpn6 is not sufficient to trigger LP2-base association. Untagged LP2, LP2, lid, or the indicated forms of Rpn6 truncation mutant LP2 (1 μ M each) were added to 1 μ M base and 10 μ M Rpn10 before analysis by native PAGE and immunoblotting. Asterisk, cross-reactive band. See also Figure S4.

diates a lid subcomplex with extra-proteasomal functions (Hofmann et al., 2009; Rinaldi et al., 2004, 2008), but we speculate that Rpn11 DUB may be held inactive in the context of LP2 and lid.

2011). The mechanism underlying this observation had been unknown. Our combined QCLMS, EM, and biochemical reconstitution studies now show that both large-scale conformational rearrangements and local repositioning of lid subunits are triggered by Rpn12 binding. These rearrangements underlie the coupling between completion of lid assembly and licensing of the lid for attachment to the base. Our findings rationalize previous genetic data implicating the Rpn12 C-terminal helix in RP formation and reveal how Rpn12 can govern lid-base joining despite making minimal contacts with the base in the mature proteasome.

Local and Global Conformational Changes during Lid Assembly

Mapping the quantified cross-links onto the 26S proteasome pseudoatomic structure reveals that the Rpn8/Rpn11 module is repositioned as assembly progresses from LP2 to lid and finally to the full proteasome (Figures 2D and 3). Recently, it has been shown that the isolated recombinant Rpn8/Rpn11 heterodimer has significant deubiquitinating (DUB) activity (Worden et al., 2014). Intriguingly, we did not detect DUB activity in LP2 (Figure S4E and data not shown) or in the full lid (data not shown). It is unclear whether the separate Rpn8-Rpn11 module occurs physiologically either as a lid assembly interme-

The repositioning suggested by our QCLMS data may be required for Rpn11 activation during incorporation into the proteasome holoenzyme, perhaps by relief of inhibitory contacts with other lid subunits. It is unlikely that Rpn6 would serve this autoinhibitory role given that Rpn6-N is extended away from the Rpn8/Rpn11 module in the EM structure of the lid. However, the high flexibility of Rpn5 (Figure 3) makes it an attractive candidate for this putative function. Notably, the Rpn5 paralog of the lid-related COP9/signalosome (CSN), CSN4, engages and inhibits the activity of the CSN6-CSN5 module (paralogous to Rpn8/Rpn11) in the fully assembled CSN (Lingaraju et al., 2014). This inhibition is released upon substrate engagement to stimulate the deneddylase activity of CSN5.

The lid is generally thought to interact stably with the base (and perhaps the CP), and Rpn11 DUB activity is regulated in part through ATP binding and/or hydrolysis by the base (Liu et al., 2006; Matyskiela et al., 2013; Verma et al., 2002). However, Rpn8 and Rpn11 make very little contact with base subunits in the proteasome holoenzyme. Our observation that Rpn12 binding to the lid at a remote location can reposition the Rpn8-Rpn11 module suggests the possibility that ATP binding or hydrolysis by the base might also allosterically regulate this module through movements of the lid rather than direct contacts with the ATPase heterohexamer itself.

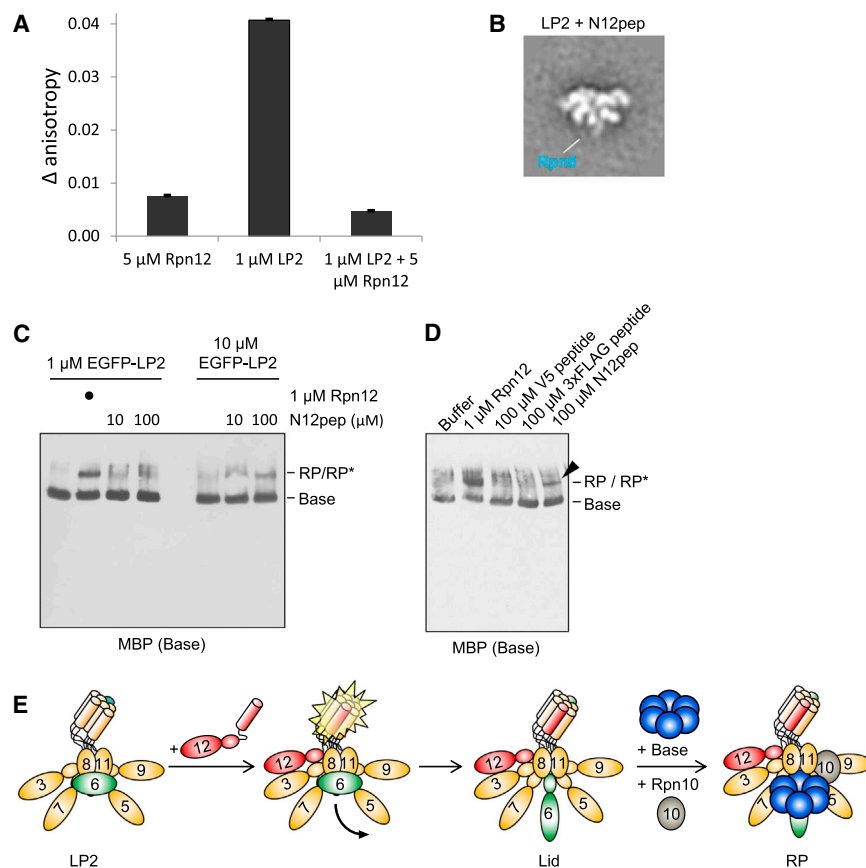


Figure 7. Engagement of the Lid Helical Bundle by the Rpn12 C-Terminal Helix Licenses the Lid for RP Assembly

(A) A fluorescently labeled peptide corresponding to amino acids 254–272 from yeast Rpn12 (N12pep) binds to LP2 similarly to full-length Rpn12. The average difference in fluorescence anisotropies between the free fluorescein-labeled N12pep and N12pep after addition of the indicated proteins are shown. Error bars, SEM.

(B) Addition of N12pep triggers extension of the Rpn6 N terminus from the body of LP2. N12pep (10 μ M) was added to purified, untagged LP2, followed by analysis by negative-stain EM. The width of the box is ~ 430 Å.

(C) N12pep triggers association of LP2 and base in the absence of Rpn12. EGFP-LP2 (1 or 10 μ M) was incubated with 1 μ M base, 10 μ M Rpn10, and the indicated concentration of N12pep before native PAGE and immunoblotting. RP*, an RP-like species devoid of Rpn12 but reactive with both lid (not shown) and base antibodies.

(D) N12pep, but not V5 or 3xFLAG control peptides, trigger LP2-BaseP association. LP2 harboring an N-terminal EGFP tag on Rpn6, base harboring an N-terminal MBP tag on Rpt1, and Rpn10 were incubated as in (C) with 100 μ M of the indicated peptides, followed by analysis as in (C). An arrowhead marks RP*.

(E) Model for Rpn12-dependent licensing of the lid for lid-base joining. LP2 assumes a closed conformation with the N-terminal domain of Rpn6 folded inward, causing steric clash with the base. During incorporation of Rpn12 into LP2, engagement of the C-terminal helical bundle triggers extension of the Rpn6 N terminus, as well as additional, as-yet-unknown events (not depicted) that promote a conformation of the lid that is competent for binding the base.

The Rpn12 C-Terminal Helix as an Allosteric Trigger

Our EM analyses and reconstitution experiments demonstrate that the Rpn12 C-terminal helix initiates a global remodeling of the lid structure from an autoinhibited conformation to one that can join to the base. Importantly, this single helix is sufficient to promote both the remodeling of LP2, as well as interaction between LP2 and base. Although the peptide was not as potent as full-length Rpn12 in the base-joining assay, this could simply reflect a lower affinity of the isolated peptide for LP2 compared to Rpn12. Other elements of Rpn12 may enhance binding of the C-terminal helix to LP2 via Rpn12 interactions with Rpn3. Additional contacts between Rpn12 and LP2 (and base) may also be necessary to promote an optimal conformation of the lid for binding to the base.

We found that Rpn6-N undergoes a striking movement from a position pressed against other subunits in LP2 to an extended conformation in the lid. Computational docking of the LP2 model onto the 26S proteasome indicates that steric clash would occur between the base ATPase ring and Rpn6-N. As neither displacement nor removal of the putative autoinhibitory domain alone was sufficient to trigger stable interaction between LP2 and the base, additional changes must also be needed. Other large steric clashes were not readily apparent

in our modeling of LP2 onto the 26S proteasome, but this could simply reflect differences in the conformation of the assembly chaperone-bound base precursor used in our study and the chaperone-free base in the context of the 26S holoenzyme observed in previous EM studies. No structural information currently exists on the chaperone-bound base precursor. The combination of QCLMS and EM approaches used here could be valuable in exploring the structure of the base precursor relative to the structure of the base in the context of the mature proteasome; such differences might also be important for regulating lid-base joining.

In this study, QCLMS was applied for the first time to study conformational changes involved in macromolecular assembly. Compared to a previous use of isotope-labeled cross-linkers for quantification (Schmidt et al., 2013), our workflow advanced a number of critical aspects: replicate analysis with label swap ensured accurate quantitation; summarizing quantitation into residue pairs reduced the impact of unrelated influences that would be felt at the peptide level (such as additional modifications or digestion efficiency); and use of well-established quantitative proteomics software minimized errors in manual analysis (Schmidt et al., 2013) and significantly reduced analysis time. These advances in turn enabled the depth and magnitude

needed for the analysis of large protein assemblies, as presented here.

Recently, a 5-subunit subcomplex of the lid (without Rpn12) was shown to bind to a chaperone-free base-CP complex *in vitro* (Yu et al., 2015). It is not yet clear whether this species points to an alternative route to RP formation, as it has not been reported to form naturally *in vivo* and it is unknown if it is competent for assembly into full 26S proteasomes. However, alternative routes to RP formation are not ruled out by our data. For example, the full lid bound to a subset of base subunits was isolated from bovine erythrocytes and shown to be capable of assembly into full RPs (Thompson et al., 2009). Alternatively, RPs or 26S proteasomes might be partially disassembled under some conditions (Fane and Prevelige, 2003; Nanduri et al., 2015).

As engagement of the LP2 helical bundle by the Rpn12 C-terminal helix was sufficient to promote interaction between LP2 and base independent of the rest of Rpn12, an allosteric signal is likely transmitted through the helical bundle. Rpn12 binding may reposition this bundle in a manner conducive for base binding. Further structural studies will be necessary to identify the specific subunit reconfigurations that drive lid-base joining. Another important challenge will be to determine the allosteric path of signal transduction from the Rpn12 C-terminal helix to the rest of the lid during conformational switching.

The Rpn12-Induced Conformational Rearrangement Is Reminiscent of Viral Maturation

Several parallels can be drawn between proteasome biogenesis and the assembly of viruses. As with the proteasomal CP and RP base, assembly of viral capsids frequently depends on exogenous chaperone proteins to form functional structures with the proper subunit arrangements (Fane and Prevelige, 2003). Further, in analogy to lid assembly and lid-base joining, incorporation of major viral subcomplexes also follows hierarchical assembly orders. For example, the head, base, and tail fibers of the T4 bacteriophage each assemble independently prior to joining together (Aksyuk and Rossmann, 2011). Finally, the capsids of many class I viruses undergo dramatic conformational rearrangements upon completion of distinct assembly events (Mateu, 2013). These rearrangements are coupled to the assembly state either via proteolytic cleavage reactions or by allosteric contacts analogous to those observed for LP2 and Rpn12. They in turn license the capsid for subsequent assembly steps, such as viral DNA packaging or attachment to the tail. Thus, while the biochemical mechanisms differ, these functional themes are shared between highly disparate macromolecular complexes and likely contribute to the faithful biogenesis of other multisubunit assemblies.

Our multipronged approach to identifying structural differences between assembly intermediates should prove valuable in future efforts to understand conformational switches and their role in proteasome assembly. The ability to reconstitute the RP *in vitro* from fully recombinant subunits will greatly facilitate biochemical and biophysical analyses of these processes. This approach should be readily extendable to studies of the assembly and action of other large multisubunit complexes.

EXPERIMENTAL PROCEDURES

Plasmids

Plasmids used in this study are listed in Table S2. Generation of multigene operons for bacterial expression was done as described (Tomko and Hochstrasser, 2014). All mutations were first introduced into host plasmids containing individual subunit expression cassettes, confirmed by DNA sequencing, and then transferred by subcloning into the appropriate operons.

Recombinant Protein Expression and Purification

All purified proteins were derived from the *S. cerevisiae* coding sequences and were purified either from *S. cerevisiae* or from *E. coli* as indicated in the text. All proteins and complexes were affinity purified using conventional techniques, followed by gel filtration. See also Supplemental Experimental Procedures.

Fluorescence Anisotropy Measurements

A peptide with residues 254–272 of Rpn12 followed by a 5,6-carboxyfluorescein-labeled lysine (NH₂-DQKTNIEKAMDYAIENIK⁺-CO₂H, herein referred to as N12pep) was synthesized by Elim Biopharmaceuticals. Fluorescent peptide (1 μM) was incubated in Lid Buffer alone, with 1 μM LP2, 5 μM Rpn12, or both at 25°C for 20 min. Fluorescence polarization was then measured on a two-channel fluorometer (PTI) with linear polarizers with excitation at 485 nm. The change in anisotropy compared to peptide alone was calculated using the average of the anisotropy values over the emission range 513–528 nm.

Assembly Assays

LP2 and its derivatives, lid, recombinant base precursor, and recombinant Rpn12 were diluted together to 1 μM each along with 10 μM Rpn10 in 26S buffer (50 mM Tris•HCl [pH 7.5], 5 mM MgCl₂, 10% glycerol, 1 mM ATP) unless otherwise noted. After incubation at 30°C for 20 min, aliquots of each reaction were separated by native polyacrylamide gel electrophoresis (native PAGE) at 100 V at 4°C as described previously (Tomko and Hochstrasser, 2014). Native PAGE-separated proteins were transferred to PVDF membranes and subjected to immunoblot analysis.

Negative-Stain Electron Microscopy

LP2 and Lid complexes were diluted to ~50 nM in Buffer A (50 mM Tris•HCl [pH 7.5], 150 mM NaCl, 5 mM MgCl₂, 10% glycerol) immediately prior to applying the sample to glow-discharged 400 mesh continuous carbon grids. After staining with 2% (w/v) uranyl acetate, the residual stain was blotted off, and the samples were air-dried in a fume hood. Data were acquired using a Tecnai F20 Twin transmission electron microscope at 120 keV at a nominal magnification of 80,000× (1.45 Å at the specimen level) using low-dose exposures (~20 e⁻ Å⁻²) with a randomly set defocus ranging from -0.5 to -1.3 μm. A total of 150–600 images of each LP2/lid complex sample were automatically recorded on a Gatan 4k × 4k CCD camera using the MSI-Raster application within the automated macromolecular microscopy software LEGION (Suloway et al., 2005). Specifically, 600, 510, 150, 150, and 300 micrographs were acquired of LP2, Lid, LP2 supplemented with purified Rpn12, LP2 with EGFP-Rpn6 complexes, and LP2 with peptide, respectively. See also Supplemental Experimental Procedures.

Cross-linking and Sample Preparation for Quantitative Cross-linking/Mass Spectrometry

LP2 or lid (100 μg) purified from yeast in Buffer HA (50 mM HEPES•NaOH, 150 mM NaCl, 5 mM MgCl₂, 10% glycerol) were each cross-linked with bis[sulfosuccinimidyl] suberate-d0 (BS³-d0) and its deuterated form bis[sulfosuccinimidyl] 2,2,7,7-tetradeuterio-suberate-d4 (BS³-d4) (Thermo Fisher) at 1:200 protein to cross-linker. The reaction was incubated at 30°C for 1 hr and quenched with 2 M glycine (pH 7.0) for 30 min at room temperature. Monomeric cross-linked lid and LP2 were isolated via SDS-PAGE and in-gel digested using trypsin (Chen et al., 2010). After digestion, cross-linked lid and LP2 samples were mixed in 1:1 molar ratio, yielding two samples for quantitative analysis: lid+BS³-d0/LP2+BS³-d4 and lid+BS³-d4/LP2+BS³-d0. Peptides were fractionated by strong cation exchange chromatography and

desalted using C18-StageTips for LC-MS/MS (Chen et al., 2010). See also Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the LP2 EM density map reported in this paper is EMD: 3136.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, two tables, and one data file and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.022>.

AUTHOR CONTRIBUTIONS

R.J.T. Jr. and M.H. conceived all experiments and wrote the manuscript with feedback from all authors. R.J.T. Jr. performed all protein purifications, cross-linking, assembly assays, and fluorescence anisotropy experiments. Z.A.C. and J.R. performed all mass spectrometry and associated data analysis. D.W.T. performed all EM analyses with advice from H.-W.W.

ACKNOWLEDGMENTS

We thank Rui Zhang for helpful comments, Eva Nogales for use of her electron microscopy resources, Andreas Martin for base expression plasmids, and Dan Finley for antibodies. This work was supported by NIH grant GM083050 to M.H. and the Wellcome Trust (Senior Research Fellowship 103139 and Centre core grant 092076) to J.R. R.J.T. Jr. was supported in part by an American Cancer Society postdoctoral fellowship. D.W.T. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG- 2218-15).

Received: April 3, 2015

Revised: June 19, 2015

Accepted: August 19, 2015

Published: October 8, 2015

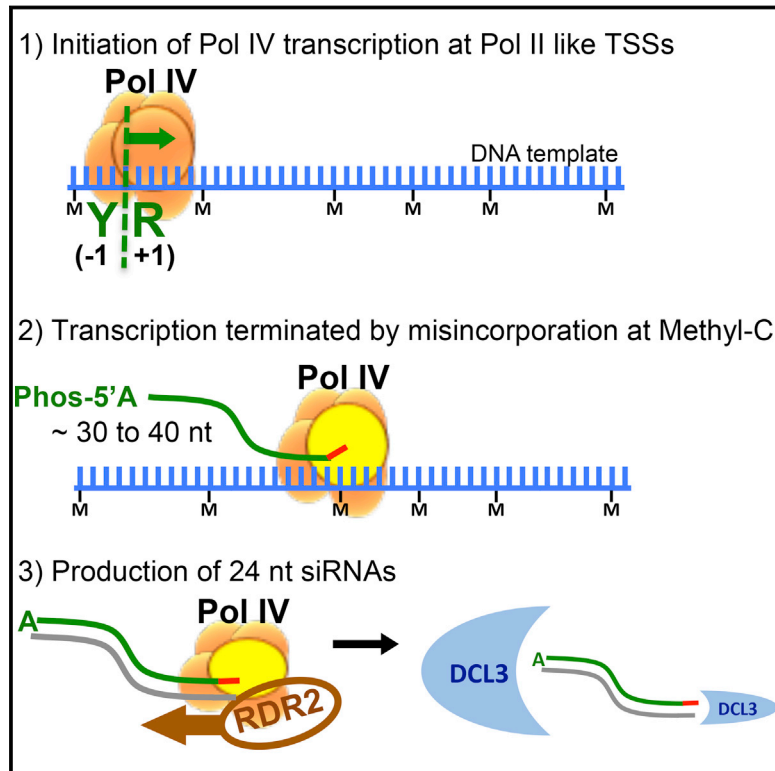
REFERENCES

- Aksyuk, A.A., and Rossmann, M.G. (2011). Bacteriophage assembly. *Viruses* 3, 172–203.
- Beckwith, R., Estrin, E., Worden, E.J., and Martin, A. (2013). Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. *Nat. Struct. Mol. Biol.* 20, 1164–1172.
- Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kaminski, T., Rasmussen, M., Larivière, L., Bukowski-Wills, J.C., Nilges, M., et al. (2010). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* 29, 717–726.
- Cheriyath, V., Jacobs, B.S., and Hussein, M.A. (2007). Proteasome inhibitors in the clinical setting: benefits and strategies to overcome multiple myeloma resistance to proteasome inhibitors. *Drugs R D* 8, 1–12.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372.
- Crawford, L.J., and Irvine, A.E. (2013). Targeting the ubiquitin proteasome system in haematological malignancies. *Blood Rev.* 27, 297–304.
- Dou, Q.P., and Zonder, J.A. (2014). Overview of proteasome inhibitor-based anti-cancer therapies: perspective on bortezomib and second generation proteasome inhibitors versus future generation inhibitors of ubiquitin-proteasome system. *Curr. Cancer Drug Targets* 14, 517–536.
- Estrin, E., Lopez-Blanco, J.R., Chacón, P., and Martin, A. (2013). Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* 21, 1624–1635.
- Fane, B.A., and Prevelige, P.E., Jr. (2003). Mechanism of scaffolding-assisted viral assembly. *Adv. Protein Chem.* 64, 259–299.
- Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* 78, 477–513.
- Fischer, L., Chen, Z.A., and Rappsilber, J. (2013). Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J. Proteomics* 88, 120–128.
- Fukunaga, K., Kudo, T., Toh-e, A., Tanaka, K., and Saeki, Y. (2010). Dissection of the assembly pathway of the proteasome lid in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* 396, 1048–1053.
- Glickman, M.H., Rubin, D.M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V.A., and Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* 94, 615–623.
- Hofmann, L., Saunier, R., Cossard, R., Esposito, M., Rinaldi, T., and Delahodde, A. (2009). A nonproteolytic proteasome activity controls organelle fission in yeast. *J. Cell Sci.* 122, 3673–3683.
- Kock, M., Nunes, M.M., Hemann, M., Kube, S., Dohmen, R.J., Herzog, F., Ramos, P.C., and Wendler, P. (2015). Proteasome assembly from 15S precursors involves major conformational changes and recycling of the Pba1-Pba2 chaperone. *Nat. Commun.* 6, 6123.
- Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., and Martin, A. (2012). Complete subunit architecture of the proteasome regulatory particle. *Nature* 482, 186–191.
- Lingaraju, G.M., Bunker, R.D., Cavadin, S., Hess, D., Hassiepen, U., Renatus, M., Fischer, E.S., and Thomä, N.H. (2014). Crystal structure of the human COP9 signalosome. *Nature* 512, 161–165.
- Liu, C.W., Li, X., Thompson, D., Wooding, K., Chang, T.L., Tang, Z., Yu, H., Thomas, P.J., and DeMartino, G.N. (2006). ATP binding and ATP hydrolysis play distinct roles in the function of 26S proteasome. *Mol. Cell* 24, 39–50.
- Mateu, M.G. (2013). Assembly, stability and dynamics of virus capsids. *Arch. Biochem. Biophys.* 531, 65–79.
- Matyskiela, M.E., Lander, G.C., and Martin, A. (2013). Conformational switching of the 26S proteasome enables substrate degradation. *Nat. Struct. Mol. Biol.* 20, 781–788.
- Nanduri, P., Hao, R., Fitzpatrick, T., and Yao, T.P. (2015). Chaperone-mediated 26S proteasome remodeling facilitates free K63 ubiquitin chain production and aggresome clearance. *J. Biol. Chem.* 290, 9455–9464.
- Nyquist, K., and Martin, A. (2014). Marching to the beat of the ring: polypeptide translocation by AAA+ proteases. *Trends Biochem. Sci.* 39, 53–60.
- Pathare, G.R., Nagy, I., Bohn, S., Unverdorben, P., Hubert, A., Körner, R., Nickell, S., Lasker, K., Sali, A., Tamura, T., et al. (2012). The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proc. Natl. Acad. Sci. USA* 109, 149–154.
- Ravid, T., and Hochstrasser, M. (2008). Diversity of degradation signals in the ubiquitin-proteasome system. *Nat. Rev. Mol. Cell Biol.* 9, 679–690.
- Richardson, P.G., Barlogie, B., Berenson, J., Singhal, S., Jagannath, S., Irwin, D., Rajkumar, S.V., Srkalovic, G., Alsina, M., Alexanian, R., et al. (2003). A phase 2 study of bortezomib in relapsed, refractory myeloma. *N. Engl. J. Med.* 348, 2609–2617.
- Rinaldi, T., Pick, E., Gambadoro, A., Zilli, S., Maytal-Kivity, V., Frontali, L., and Glickman, M.H. (2004). Participation of the proteasomal lid subunit Rpn11 in mitochondrial morphology and function is mapped to a distinct C-terminal domain. *Biochem. J.* 381, 275–285.
- Rinaldi, T., Hofmann, L., Gambadoro, A., Cossard, R., Livnat-Levanon, N., Glickman, M.H., Frontali, L., and Delahodde, A. (2008). Dissection of the carboxyl-terminal domain of the proteasomal subunit Rpn11 in maintenance of mitochondrial structure and function. *Mol. Biol. Cell* 19, 1022–1031.
- Schmidt, C., Zhou, M., Marriott, H., Morgner, N., Politis, A., and Robinson, C.V. (2013). Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation. *Nat. Commun.* 4, 1985.
- Sharon, M., Taverner, T., Ambroggio, X.I., Deshaies, R.J., and Robinson, C.V. (2006). Structural organization of the 19S proteasome lid: insights from MS of intact complexes. *PLoS Biol.* 4, e267.

- Sharon, M., Witt, S., Glasmacher, E., Baumeister, W., and Robinson, C.V. (2007). Mass spectrometry reveals the missing links in the assembly pathway of the bacterial 20 S proteasome. *J. Biol. Chem.* **282**, 18448–18457.
- Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* **151**, 41–60.
- Thompson, D., Hakala, K., and DeMartino, G.N. (2009). Subcomplexes of PA700, the 19 S regulator of the 26 S proteasome, reveal relative roles of AAA subunits in 26 S proteasome assembly and activation and ATPase activity. *J. Biol. Chem.* **284**, 24891–24903.
- Tomko, R.J., Jr., and Hochstrasser, M. (2011). Incorporation of the Rpn12 subunit couples completion of proteasome regulatory particle lid assembly to lid-base joining. *Mol. Cell* **44**, 907–917.
- Tomko, R.J., Jr., and Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. *Annu. Rev. Biochem.* **82**, 415–445.
- Tomko, R.J., Jr., and Hochstrasser, M. (2014). The intrinsically disordered Sem1 protein functions as a molecular tether during proteasome lid biogenesis. *Mol. Cell* **53**, 433–443.
- Unverdorben, P., Beck, F., Śledź, P., Schweitzer, A., Pfeifer, G., Plitzko, J.M., Baumeister, W., and Förster, F. (2014). Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc. Natl. Acad. Sci. USA* **111**, 5544–5549.
- Verma, R., Aravind, L., Oania, R., McDonald, W.H., Yates, J.R., 3rd, Koonin, E.V., and Deshaies, R.J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **298**, 611–615.
- Witt, S., Kwon, Y.D., Sharon, M., Felderer, K., Beuttler, M., Robinson, C.V., Baumeister, W., and Jap, B.K. (2006). Proteasome assembly triggers a switch required for active-site maturation. *Structure* **14**, 1179–1188.
- Worden, E.J., Padovani, C., and Martin, A. (2014). Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. *Nat. Struct. Mol. Biol.* **21**, 220–227.
- Yu, Z., Livnat-Levanon, N., Kleifeld, O., Mansour, W., Nakasone, M.A., Castaneda, C.A., Dixon, E.K., Fushman, D., Reis, N., Pick, E., and Glickman, M.H. (2015). Base-CP proteasome can serve as a platform for stepwise lid formation. *Biosci. Rep.* **35**, e00194.

A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis

Graphical Abstract



Authors

Jixian Zhai, Sylvain Bischof, Haifeng Wang, ..., Blake C. Meyers, Israel Ausin, Steven E. Jacobsen

Correspondence

israelausin@fafu.edu.cn (I.A.), jacobsen@ucla.edu (S.E.J.)

In Brief

In contrast to the assumption that one Pol IV transcript gives rise to multiple siRNAs, it is now shown that Pol IV transcribes 30- to 40-nt short RNAs that are further processed to generate 24-nt mature products, thereby demonstrating a “one precursor, one siRNA” model for Pol IV-dependent siRNA biogenesis in Arabidopsis.

Highlights

- Pol IV transcribes 30 to 40 nucleotide short RNAs to initiate siRNA biogenesis
- P4RNAs mirror siRNAs in distribution, abundance, strand bias and 5'-A preference
- P4RNAs feature TSSs similar to Pol II products
- P4RNAs carry 3'-misincorporation that preferentially occurs at methylated cytosines

Accession Numbers

GSE61439



A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis

Jixian Zhai,^{1,8} Sylvain Bischof,^{1,8} Haifeng Wang,^{2,1} Suhua Feng,¹ Tzuu-fen Lee,³ Chong Teng,³ Xinyuan Chen,⁴ Soo Young Park,⁵ Linshan Liu,⁵ Javier Gallego-Bartolome,^{1,6} Wanlu Liu,¹ Ian R. Henderson,^{1,9} Blake C. Meyers,³ Israel Ausin,^{2,*} and Steven E. Jacobsen^{1,7,*}

¹Department of Molecular, Cell and Developmental Biology, University of California at Los Angeles, Los Angeles, CA 90095, USA

²Haixia Institute of Science and Technology (HIST), Fujian Agriculture and Forestry University, Fuzhou 350002, China

³Department of Plant and Soil Sciences and Delaware Biotechnology Institute, University of Delaware, Newark, DE 19716, USA

⁴Department of Microbiology, Immunology, and Molecular Genetics, University of California at Los Angeles, Los Angeles, CA 90095, USA

⁵Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, CA 90095, USA

⁶Plant Biology Laboratory and Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

⁷Howard Hughes Medical Institute, University of California at Los Angeles, Los Angeles, CA 90095, USA

⁸Co-first author

⁹Present address: Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK

*Correspondence: israelausin@fafu.edu.cn (I.A.), jacobsen@ucla.edu (S.E.J.)

<http://dx.doi.org/10.1016/j.cell.2015.09.032>

SUMMARY

RNA-directed DNA methylation in *Arabidopsis thaliana* is driven by the plant-specific RNA Polymerase IV (Pol IV). It has been assumed that a Pol IV transcript can give rise to multiple 24-nt small interfering RNAs (siRNAs) that target DNA methylation. Here, we demonstrate that Pol IV-dependent RNAs (P4RNAs) from wild-type *Arabidopsis* are surprisingly short in length (30 to 40 nt) and mirror 24-nt siRNAs in distribution, abundance, strand bias, and 5'-adenine preference. P4RNAs exhibit transcription start sites similar to Pol II products and are featured with 5'-monophosphates and 3'-misincorporated nucleotides. The 3'-misincorporation preferentially occurs at methylated cytosines on the template DNA strand, suggesting a co-transcriptional feedback to siRNA biogenesis by DNA methylation to reinforce silencing locally. These results highlight an unusual mechanism of Pol IV transcription and suggest a “one precursor, one siRNA” model for the biogenesis of 24-nt siRNAs in *Arabidopsis*.

INTRODUCTION

In *Arabidopsis*, 24-nt siRNAs are the triggers for RNA-directed DNA methylation (RdDM), which plays central roles in repressing transposable elements (TEs) and maintaining genome integrity (Law and Jacobsen, 2010; Matzke and Mosher, 2014). The current model for 24-nt siRNA biogenesis is composed of several sequential steps. First, Pol IV recognizes heterochromatic regions, in part via SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) (Law et al., 2013), and transcribes precursor RNAs. These precursor RNAs are then thought to be processed by RNA-dependent RNA polymerase 2 (RDR2)

to form double-stranded RNAs (dsRNAs). The dsRNAs are primarily cleaved by dicer-like 3 (DCL3) to produce 24-nt siRNAs (Matzke and Mosher, 2014). Although Pol IV is clearly required for the biogenesis of 24-nt siRNAs (Herr et al., 2005; Kanno et al., 2005; Onodera et al., 2005; Pontier et al., 2005) and was shown to be transcriptionally active in vitro (Haag et al., 2012), the nature and characteristics of Pol IV-dependent transcripts remain poorly understood. The low transcription level at silent Pol IV loci and the efficient downstream processing by dicers has made Pol IV transcripts difficult to study. Recently, regions containing Pol IV-dependent transcripts were described in a *dcl2/3/4* triple mutant that compromises downstream processing of siRNAs (Li et al., 2015). In that study, with the conventional assumption that Pol IV transcribes long precursors, RNA was fragmented before cloning, and sequenced reads were assembled into contiguous regions that corresponded to well-known siRNA-producing regions (Li et al., 2015). In the present study, we sought to identify the very low abundance Pol IV-dependent transcripts in a wild-type background, and to precisely characterize their start and end positions. We utilized a special method for RNA sequencing named PATH (parallel analysis of tail and head) that utilizes RNA adapters to capture both ends of any RNA greater than 27-nt with a 5' monophosphate and a 3' hydroxyl. Using PATH, we efficiently cloned Pol IV-dependent RNAs (P4RNAs) and found that they were only on the order of 30 to 40 nt in length. These P4RNAs mirror 24-nt Pol IV-dependent siRNAs in distribution, abundance, strand bias, and 5'-adenine preference, suggesting that they are direct precursors that often determine both the start position and strandedness of siRNAs. We also observed extensive incorporation of 3'-non-templated nucleotides preferentially at methylated cytosines on the template DNA strand, suggesting that Pol IV misincorporates and terminates at DNA methylated sites. Our results support a “one precursor, one siRNA” model for the biogenesis of Pol IV-dependent 24-nt siRNAs in *Arabidopsis* that creates a positive feedback loop between siRNA biogenesis and DNA methylation.

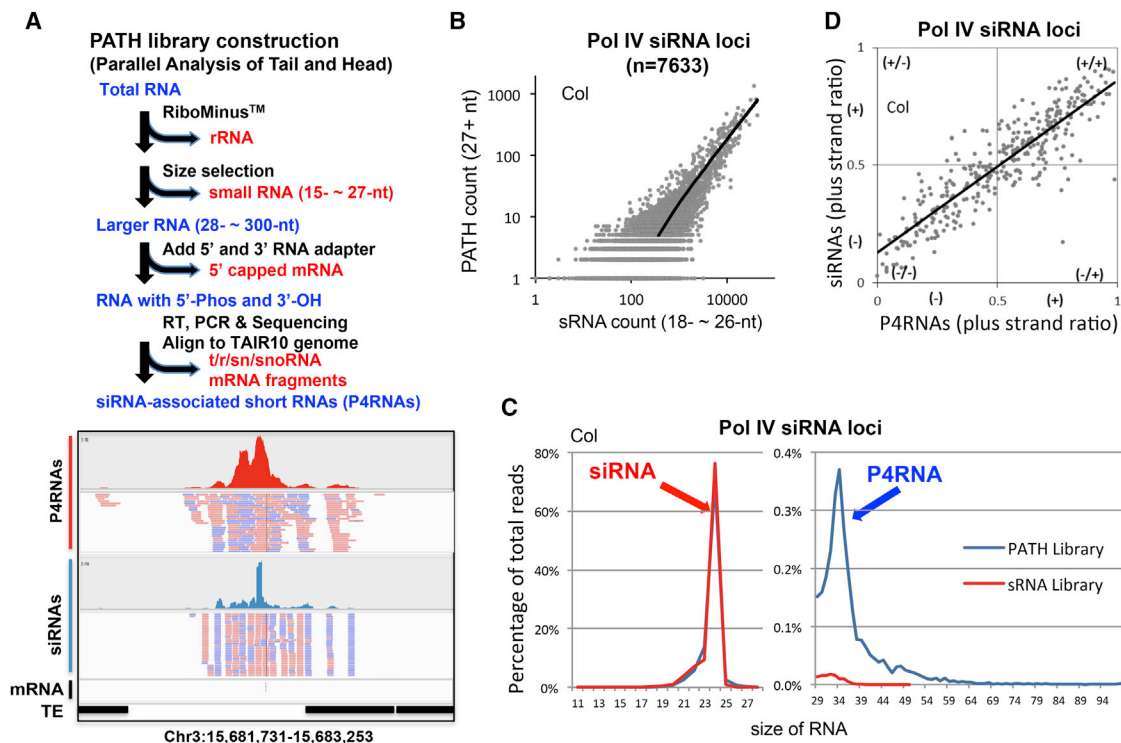


Figure 1. Identification of Short RNAs Associated with siRNA Loci

(A) Procedure for the construction and analysis of PATH libraries. An example siRNA locus is shown at the bottom with IGV screenshots of P4RNAs and siRNAs matching to that region.

(B) Abundances are highly correlated for PATH reads (27+ nt) and sRNA reads (18 to 26 nt) in wild-type Col from previously defined Pol IV siRNA loci (Law et al., 2013).

(C) Size distribution plots of all PATH reads and sRNA reads in Col from Pol IV siRNA loci. A distinct peak at 30 to 40 nt can be seen in the PATH library; we named these “P4RNAs.”

(D) P4RNAs and siRNAs share the same strand bias at Pol IV siRNA loci. Only Pol IV siRNA loci matched by more than 100 P4RNAs in Col were selected to obtain a robust calculation of strandedness. The plus-strand ratio was calculated as the abundance of reads matching to the plus strand divided by the total number of reads at that locus.

See also Figures S1, S2, and S3 and Tables S1, S2.

RESULTS

Identification of Short RNAs Derived from Pol IV-Dependent 24-nt siRNA Loci

We utilized an RNA cloning scheme called PATH that uses RNA ligation and gel-based size selection to capture RNAs larger than 27 nt that contain a 5' monophosphate and a 3' hydroxyl (Figure 1A, see Experimental Procedures). From a wild-type Columbia (Col) library, around 162 million PATH reads were obtained, of which ~50 million could be mapped uniquely to the *Arabidopsis* genome (Table S1), and PATH reads with a minimal length of 27 nt were used for further analysis. Although the majority of PATH reads matched to structural RNAs (t/r/sn/snoRNA) (Table S1), a large number of reads mapped to previously defined Pol IV siRNA loci (Law et al., 2013) (Figure 1A and Table S2). In addition to their co-localization, the abundances of siRNAs and PATH reads were also highly positively correlated (Figure 1B). At Pol IV siRNA loci, in addition to the 24-nt siRNA peak that is still present in PATH libraries (Figure 1C), we observed a secondary peak of PATH reads (P4RNAs) that pri-

marily ranged from 30 to 40 nt (Figure 1C). The size distribution of P4RNAs was different from other PATH reads in the libraries such as reads matching tRNAs and snoRNAs (Figure S1A). The short length of P4RNAs fits well with the specific preference of DCL3 for short (30 to 50 bp) dsRNA substrates (Nagano et al., 2014), consistent with the hypothesis that these short P4RNAs serve as the precursors for siRNA biogenesis.

Biogenesis and Processing of Pol IV RNAs

To investigate the biogenesis of P4RNAs, we constructed small RNA and PATH libraries from null mutants of *NRPD1* (encoding the largest subunit of Pol IV), *NRPD/E2* (encoding the shared second-largest subunit of both Pol IV and Pol V), *RDR2*, as well as from the double mutants of *DCL3 NRPD1* and *DCL3 RDR2*. We found that both siRNAs and P4RNAs were eliminated in all these mutant backgrounds, indicating a complete dependency of P4RNA biogenesis on both Pol IV and RDR2 (Figure 2A). Although the loss of P4RNAs in *rdr2* mutants is somewhat unexpected given the assumed downstream function of RDR2 (Haag et al., 2012), this result is consistent with a previous report

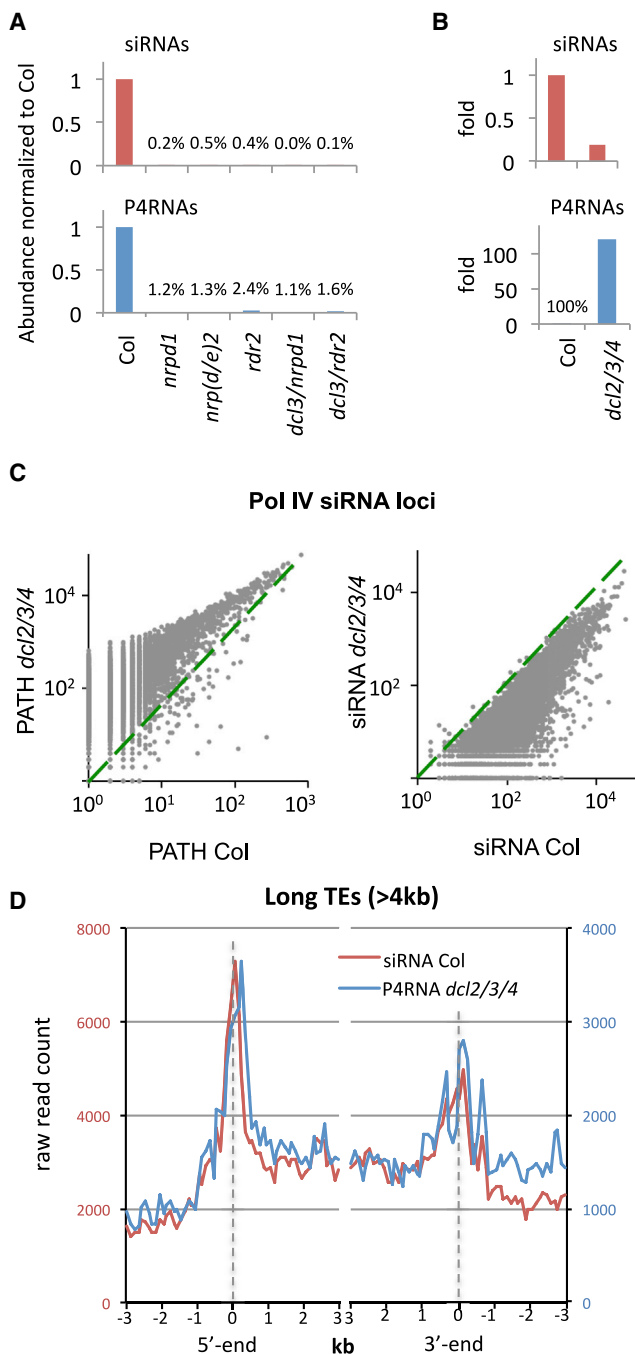


Figure 2. Biogenesis and Processing of P4RNAs

(A) Both Pol IV siRNAs and P4RNAs are eliminated in *nprp1*, *nprp(d/e)2*, *rdr2*, *dcl3/nprp1*, and *dcl3/rdr2*. Small RNA abundances were normalized to the sum of all TAIR10-annotated miRNAs in each sRNA library, and then compared to Col; P4RNA abundances were normalized to the total number of reads in each PATH library.

(B) Pol IV siRNAs are reduced in *dcl2/3/4*, whereas P4RNAs accumulate substantially.

(C) Abundances of P4RNA and siRNA at each locus in Col and *dcl2/3/4*.

(D) At long TEs, P4RNAs resemble the distribution of siRNAs – enriched at promoters and termini of TEs, and also spread into TE bodies.

See also Tables S1 and S2.

(Li et al., 2015), and suggests that in vivo, Pol IV and RDR2 activities are tightly coupled. These results are also consistent with the earlier observation that RDR2 is in tight association with Pol IV subunits in vivo (Haag et al., 2012; Law et al., 2011).

To test if P4RNA biogenesis is dependent on downstream processing events by dicer proteins, we carried out additional siRNA and PATH library analysis from plants triply mutant for *DCL2*, 3, and 4. Consistent with a recent report detecting Pol IV dependent transcripts in a *dcl2/3/4* mutant background (Li et al., 2015), we observed a dramatically increased accumulation of P4RNAs in the triple mutant *dcl2/3/4* (Figures 2B and 2C). In contrast, siRNAs levels were dramatically decreased in *dcl2/3/4* at almost all loci (Figures 2B and 2C) and their sizes were shifted from predominantly 24 nt to predominantly 21 nt (Figure S1B). In addition, we observed that the abundances of the remaining siRNAs in *dcl2/3/4* were tightly correlated with the abundances of siRNAs in wild-type (Figure 2C). We also found that the abundances of P4RNAs in *dcl2/3/4* were strongly positively correlated with the abundances of siRNAs in Col (Figure S1C). As one example, P4RNAs and siRNAs showed a very similar enrichment at the boundaries of long transposable elements (Figure 2D). The accumulation of P4RNAs but reduction of siRNAs in *dcl2/3/4* mutants are inconsistent with the interpretation that P4RNAs are merely longer, misprocessed siRNAs, and suggests that P4RNAs are indeed the precursors of Pol IV-dependent siRNAs.

We also analyzed siRNA and PATH libraries from other *DCL* mutant combinations (*dcl2/4*, *dcl3*, *dcl2/3*, and *dcl3/4*) and compared these with *dcl2/3/4*. As previously reported (Henderson et al., 2006), all *dcl3* mutant combinations showed a shift in siRNA size from 24 nt to 21 and/or 22 nt (Figure 3A, top). PATH sequencing indicated that the over-accumulation of P4RNAs was negligible in *dcl2/4*; higher in *dcl3*, *dcl2/3*, and *dcl3/4*; and highest in *dcl2/3/4*, indicating that the P4RNA processing by *DCL2* and *DCL4* is less efficient than by *DCL3* (Figure 3A, bottom). As a secondary confirmation of the accumulation of P4RNAs in *dcl3* mutants, we examined small RNA northern blots for the presence of 30- to 40-nt signals. Using LNA probes to six different siRNA producing loci we indeed observed an RNA smear with larger sizes, along with an accumulation pattern (highest in *dcl2/3/4*) that matched the trend of the P4RNA level in our PATH libraries (Figures 3B and 3C) (Henderson et al., 2006). In addition, we used a reverse transcription qPCR (real-time RT-PCR) approach to verify the very short nature of the P4RNAs. Because the peak of P4RNA abundance is 30–40 nt, whereas longer P4RNAs are much less abundant (Figure 1C), we reasoned that primers spaced about 40 nt from each other should amplify P4RNAs much more efficiently than those spaced slightly further apart. After normalizing for efficiency of DNA amplification, we observed a 100-fold higher level of PCR product when comparing a 36 bp versus a 58 bp amplicon at one locus, and 50-fold higher level of PCR product when comparing a 39 bp versus an 81 bp amplicon at a second locus (Figure S2A).

To directly test whether the longer P4RNAs might simply be misprocessed siRNAs that are loaded into AGO4, we performed AGO4 RNA-IP (RIP) in the *dcl2/3/4* genetic background using antibodies against endogenous AGO4 and then characterized AGO4-associated RNAs with high-throughput sequencing. The

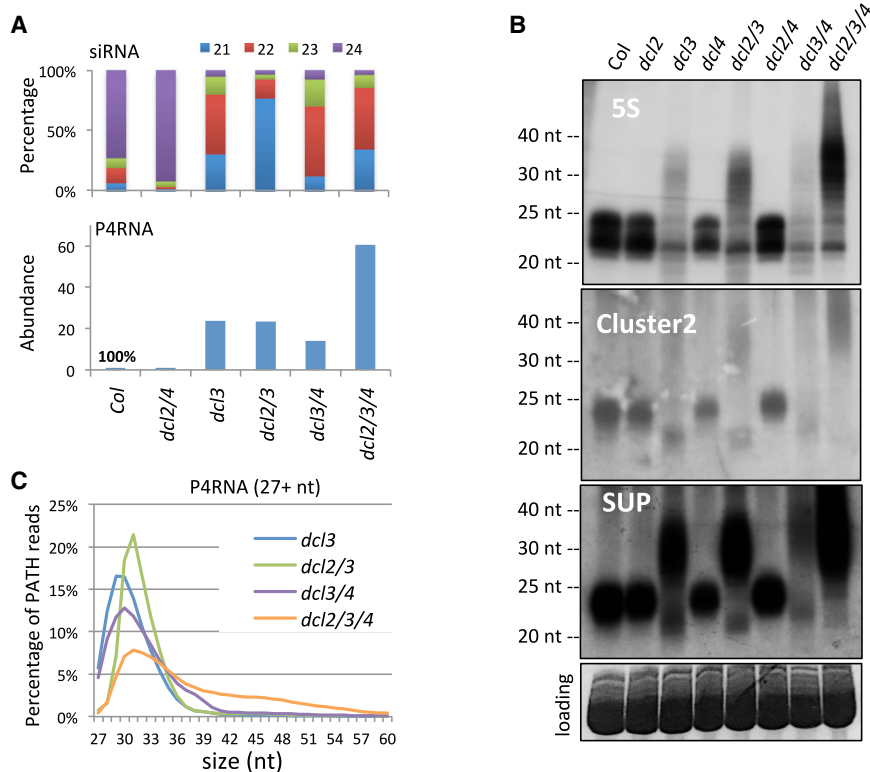


Figure 3. P4RNAs in *dcl* Mutant Combinations

(A) Size of Pol IV siRNAs shifted from 24 nt in Col to 21/22 nt in *dcl3*-containing mutants, accompanied by the accumulation of P4RNAs. P4RNA abundances are normalized to that in Col.

(B) RNA blot analysis of siRNAs and P4RNAs in *dcl* mutant backgrounds.

(C) Size distribution of P4RNAs in *dcl3*-containing mutants.

See also Figure S3 and Tables S1 and S2.

shared strandedness of P4RNAs and siRNAs, suggest that the 30 to 40 nt P4RNAs serve as precursors to Pol IV dependent siRNAs.

Pol IV Transcription Initiates at Pol II-like TSSs and Favors 5'-Adenine

Next we investigated the nature of the 5' ends of P4RNAs. Given that our cloning method only captures RNAs with 5-monophosphates (Figure 1A), we sought to measure the proportion of P4RNAs containing this type of 5' end. To address this, we used Terminator exonuclease to preferentially digest RNAs with a 5'-monophosphate. We subsequently measured the abundance of the remaining P4RNAs by real-time RT-PCR. Consistent with a recent study (Li et al., 2015), Terminator treatment degraded the majority of P4RNAs at all loci tested (Figure S2B). Thus, while it is possible that subpopulations of P4RNAs have other end structures such as 5'-triphosphates, 5'-caps, or 5'-hydroxyl groups, it appears that the majority of P4RNAs contain 5'-monophosphates.

Because our method did not include a fragmentation step that is typical in RNA-seq library protocols (Li et al., 2015), it was possible to detect the 5' nucleotide of P4RNA reads. We observed a strong enrichment of T/C (Y) at the -1 position (the nucleotide immediately upstream of the first nucleotide of the P4RNA read) and A/G (R) at the +1 position (beginning nucleotide of the read) (Figure 4A). Further, the four possible Y/R dinucleotides at the -1/+1 positions were by far the most enriched dinucleotides at the 5' end of P4RNAs (Figure 4B). This pattern is very similar to that known for the transcriptional start sites (TSSs) of RNA Polymerase II (Pol II) in plants and other organisms and it is referred to as the "Y/R rule" (Cumbie et al., 2015; Nechaev et al., 2010; Yamamoto et al., 2007). This result suggests that Pol IV has retained this preference from its evolutionary ancestor Pol II (Ream et al., 2009), and that the 5' ends of P4RNAs likely represent Pol IV transcriptional start sites. The short TSS like sequences at the 5' ends of P4RNA, along with their very short nature does not support previous models in which Pol IV initiates transcription solely at the nucleosome-depleted promoter regions near the ends of transposons to produce long transcripts (Li et al., 2015). Instead, our results suggest that Pol IV can initiate transcription at many positions that

AGO4 RIP result showed that even though there is a massive accumulation of longer P4RNAs and a reduction of 24nt siRNAs, AGO4 still selectively binds to the remaining 22-24 nt siRNAs but not the longer P4RNAs (Figure S2C), strongly supporting our model that P4RNAs are the precursors of the 24-nt siRNAs.

Even though double-stranded RNAs are thought to be intermediates of siRNA biogenesis, it is known that at many loci in the genome, siRNAs are predominantly found on one strand of the genome but not the other (Lister et al., 2008; Zhong et al., 2014). In addition, these strand-biased clusters of siRNAs correspond to strand biased DNA methylation (Lister et al., 2008; Zhong et al., 2014). At these strand-biased siRNA clusters, we also observed strongly strand-biased P4RNAs in both Col and *dcl2/3/4* (Figures 1D and S1D). These results suggest that our PATH libraries are primarily composed of single-stranded, Pol IV-derived strands, but not the RDR2-derived second strand. In support of this hypothesis, we used RT-PCR to amplify these strand-biased clusters, and indeed found that Pol IV/RDR2-dependent transcripts could be amplified equally from both strands (Figure S3). This suggests that RDR2 strands are likely present in vivo but are not cloned by the specific PATH sequencing technique employed in this study. The observation that P4RNAs and siRNAs show the same strandedness suggests that the Pol IV-derived strands, rather than the RDR2-derived strands, are strongly favored to become the final 24-nt siRNA products.

In summary, results from siRNA and PATH libraries of mutants deficient in siRNA biogenesis or processing, together with the

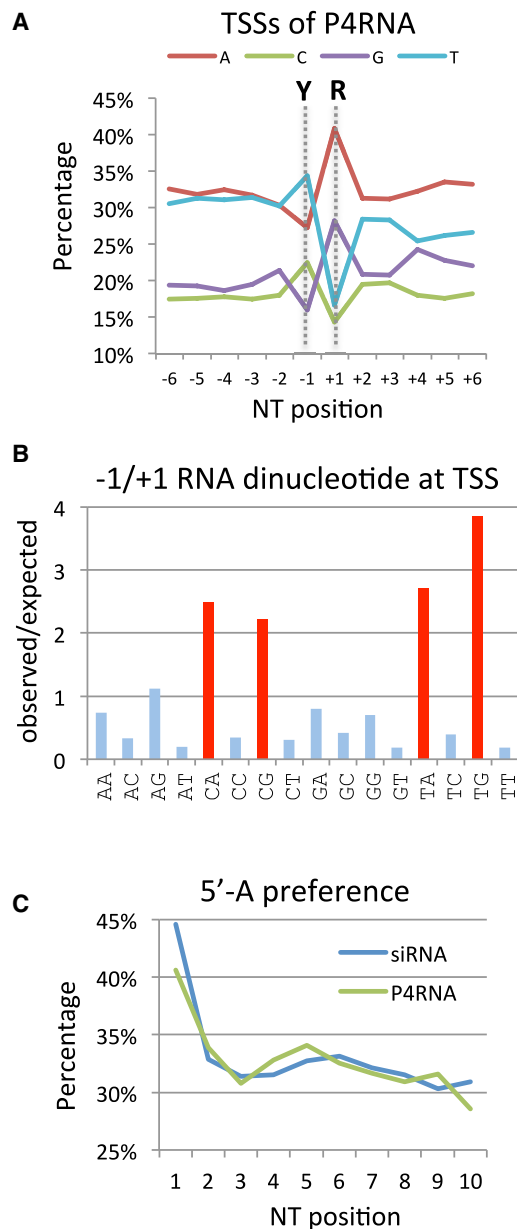


Figure 4. P4RNAs Feature Pol II-like TSSs and Favor 5'-Adenine

(A) TSSs of Pol IV exhibit preference for C/T at the -1 position and A/G at the $+1$ position, resembling the "Y/R rule" of Pol II TSS, calculated using P4RNAs from *dcl2/3/4*.

(B) Dinucleotide enrichment at $-1/+1$ of P4RNA, calculated using P4RNAs from *dcl2/3/4*.

(C) Both siRNA and P4RNAs from Col have a strong preference for 5' adenine. See also Figures S4 and S5.

resemble the Y/R features of Pol II TSSs, transcribing many short P4RNAs along the length of transposons.

Because Pol IV transcripts feature Pol II-like TSSs, we also performed genome-wide profiling of Pol II occupancy in *Arabidopsis* via ChIP-seq in wild-type and different mutant backgrounds (*nprpd1* and *dcl2/3/4*), and compared this with Pol

IV ChIP-seq (Law et al., 2013). Our results showed that Pol II does not appear to access Pol IV loci even in the absence of Pol IV (*nprpd1*) (Figure S4A). Furthermore, whole-genome bisulfite sequencing and small RNA sequencing of floral tissues from the weak Pol II mutant (*nprpd2-3*) did not reveal obvious changes in either DNA methylation or siRNA biogenesis (Figures S4B and S4C), which is consistent with our previous analysis of *nprpd2-3* using leaf tissue (Stroud et al., 2013). Therefore despite the proposed crosstalk between Pol II and the RdDM pathway (Zheng et al., 2009), our data suggest that Pol II and Pol IV occupy distinct territories on the genome.

Arabidopsis 24-nt siRNAs are primarily loaded into AGO4 and are strongly biased toward having a 5'-adenine, which was previously shown to involve an AGO4 loading preference (Havecker et al., 2010; Mi et al., 2008). Interestingly, we found that P4RNAs, like siRNAs, also show a strong enrichment for 5'-adenine (Figures 4C, S5A, and S5B). Because the size of P4RNAs is approximately 30 to 40 nt, on average only one 24-nt siRNA duplex could be processed from each of these P4RNA precursors. This fact, coupled with the shared 5' adenine preference and the shared strand preference, suggests that 24-nt siRNAs are preferentially cleaved from the 5' portion of P4RNAs. Consistent with this hypothesis, DCL3 was shown to prefer short double-stranded RNAs (30 to 50 bp) that contain a 5' adenine (Nagano et al., 2014). Therefore, our results favor a scenario in which the 5'-adenine preference of P4RNAs likely contributes to the 5'-adenine preference of Pol IV siRNAs. In addition, our results provide a plausible explanation as to why AGO4 evolved to bind siRNAs with 5' adenine. Taken together, the short size and 5'-A feature of Pol IV transcripts may help to channel their processing to DCL3 rather than dicer proteins in other silencing pathways, and thus lead to production of predominantly 24-nt siRNAs at Pol IV transcribed loci.

Pol IV Transcription Preferentially Terminates at Methylated Cytosines with Misincorporated Nucleotides

We analyzed the sequence composition of P4RNA reads with perfect match to genome and found enrichment for A, C, and U at last three positions of the 3' end (Figure S5C). It is not known whether Pol IV tends to cease transcription at this sequence, or whether it might transcribe a longer RNA that is then processed by an unknown endonuclease. We found a similar compositional bias at the 3' end of Pol IV-dependent siRNAs (Figure S5D), although the magnitude of the biases were lower for siRNAs, suggesting that P4RNAs are processed at some level at their 3' ends to produce siRNAs. Despite the ACU enrichment at the 3' end of siRNAs, they still show enrichment for 5' adenine, which is likely explained by the AGO4 preference for loading siRNAs with a 5' adenine. These results provides additional evidence that the P4RNAs described here are indeed the precursors of siRNAs, and it is again consistent with (1) the shared strandedness of P4RNAs and siRNAs, and (2) the hypothesis that the Pol IV strand, rather than the RDR2 strand, is favored as the final siRNA.

We also performed an analysis in which we allowed multiple mismatches during genome mapping of P4RNAs. Interestingly, we found that more than half of the P4RNAs contained one or

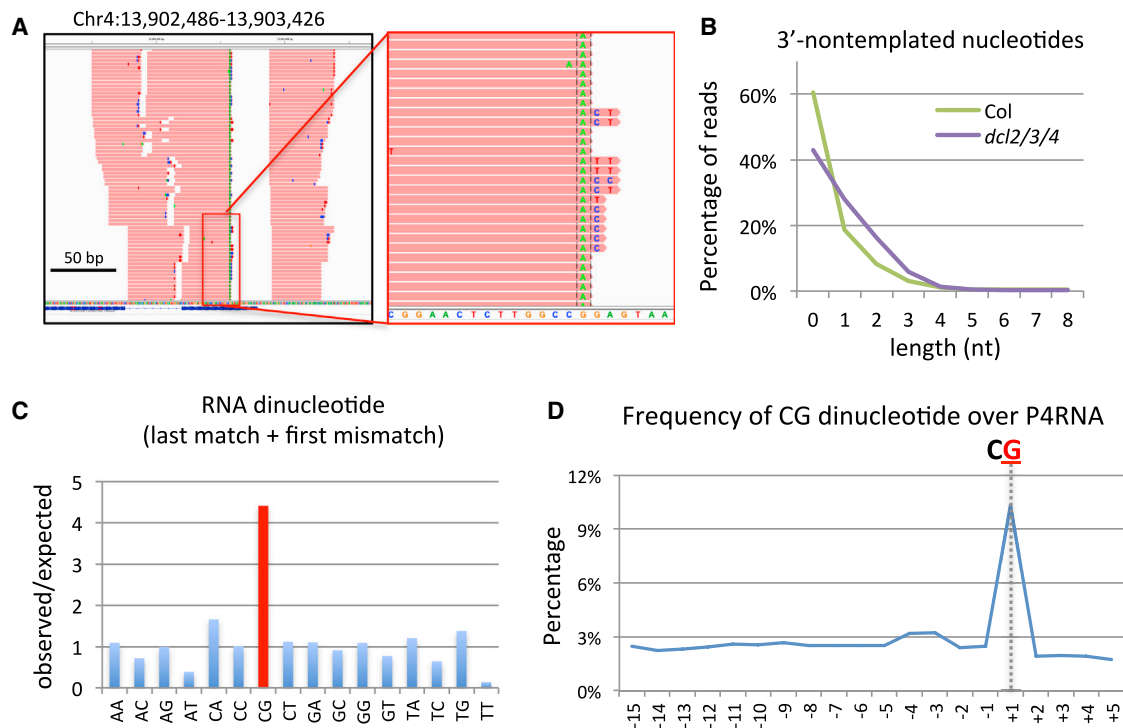


Figure 5. Pol IV Transcription Preferentially Terminates at Methylated Cytosines with Misincorporated Nucleotides

(A) Example of 3'-non-templated nucleotides on P4RNAs. The black bar represents 50 bp in length.

(B) Length of the 3'-end non-templated nucleotides, defined by the first mismatched nucleotide to the last nucleotide. If a P4RNA has no mismatch, the length is zero.

(C) Di-nucleotide enrichment at the first mismatched position at the 3' end.

(D) Frequency of CG dinucleotide on reference sequence over the P4RNAs with misincorporation, "-1" marks the last perfectly matched position, and "+1" marks the first mismatched position. The count of each CG is designated to the position of the G, therefore the peak at "+1" represents a peak of CG at "-1/+1." See also Figure S6.

two non-templated nucleotides at their 3'-ends (Figures 5A and 5B). In contrast, in the same PATH libraries, reads derived from Pol II-transcribed coding regions, or from microRNA processing intermediates, had very few mismatches, and these mismatches were not localized to the 3' ends (Figures S6A and S6B). To further rule out the possibility that P4RNA 3' non-templated nucleotides are due to lower quality of sequencing toward the end of the read or result from incomplete trimming of adaptor sequence, we analyzed the second read (read2) from the paired-end sequencing. On read2, the beginning nucleotide corresponds to the 3' end nucleotide of RNA, where base quality is high and there is no trimming step involved, and we observed the same high-level of non-templated nucleotides at the 3'-end of P4RNA (Figure S6C).

The 3' end base composition of P4RNAs containing non-templated nucleotides was quite different from those with a perfect genome match (Figures S5C and S5E), suggesting that P4RNAs with non-templated nucleotides terminate by a different mechanism than those without. On the other hand, the 5' end base composition of reads with non-templated nucleotides was very similar to those without, suggesting that both classes may share a similar transcription initiation mechanism by Pol IV (Figures S5F and S5G). All four nucleotides were found among the non-

templated nucleotides, although there was some preference for guanines (Figure S7A). In addition, we observed different preferences for non-templated nucleotides depending on the sequence that should have been present, suggesting that the preference for a particular non-templated nucleotide is determined by the sequence of the DNA template for Pol IV (Figures S7B and S7C). Because all four nucleotides were present, and because different incorrect nucleotides were present depending on the template DNA sequence, it seems most likely that these nucleotides arise from misincorporation during Pol IV transcription rather than from the activity of a terminal transferase.

Because misincorporation of nucleotides occurred most frequently at positions corresponding to guanines (which would be cytosines on the DNA template) (Figure 5A), we hypothesized that misincorporation might be caused by in part by cytosine DNA methylation. The most highly methylated sequences in the *Arabidopsis* genome are CG dinucleotides (Cokus et al., 2008; Lister et al., 2008), and Pol IV siRNA loci targeted by RdDM are usually heavily methylated at most CG sites (Stroud et al., 2013) (Figure S4B). Agreeing with the DNA methylation hypothesis, we found that CG dinucleotides (G being the last nucleotide of the RNA) exhibited by far the highest enrichment of misincorporation among the 16 possible dinucleotide

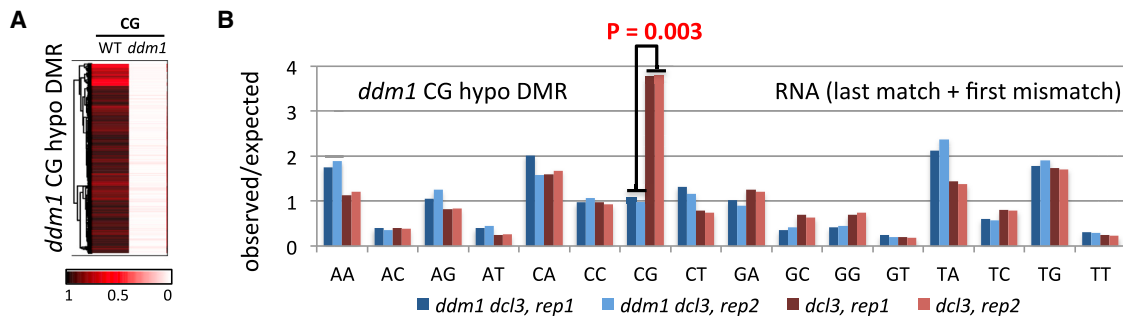


Figure 6. Pol IV Transcription Preferentially Terminates at Methylated Cytosines with Misincorporated Nucleotides

(A) CG hypomethylated DMRs in *ddm1* compared to a wild-type control.

(B) 3'-misincorporated nucleotides of P4RNA at *ddm1* CG DMRs in *ddm1 dcl3* compared to *dcl3*.

See also Figure S7.

sequences (Figure 5C). In addition, CG dinucleotides were strongly enriched at the 3'-end of P4RNAs that exhibited misincorporation (Figure 5D). The second group of most commonly methylated sequences in the genome and at Pol IV siRNA loci are CHG sites (Figure S4B) (Cokus et al., 2008; Lister et al., 2008). A trinucleotide analysis showed that all of the trinucleotides showing the strongest tendency for misincorporation were those that contained CG sites (Figure S7D). In addition, CHG sites showed a strong tendency for misincorporation, which was higher than AHG, THG, or GHG sites (Figure S7D).

To directly test whether the loss of methylation can alter the pattern of 3' end misincorporation, we analyzed the *ddm1* (*decrease in DNA methylation 1*) mutant that exhibits a severe loss of methylation in heterochromatin (Matzke and Mosher, 2014). Despite the significant loss of methylation, many silent loci are still producing 24-nt siRNAs in *ddm1* (Colomé-Tatché et al., 2012), suggesting that Pol IV is still largely functional in *ddm1*. Because P4RNAs are elevated in *dcl3* mutants, we utilized a small RNA dataset from *ddm1 dcl3* double mutant (RNAs smaller than 200 nt) (McCue et al., 2015) and compared these with a similar dataset from the *dcl3* single mutant. We first focused our analysis on a set of strong *ddm1* hypomethylated CG DMRs (differentially methylated regions) (Figure 6A, Experimental Procedures) from whole-genome bisulfite sequencing data of *ddm1* (Creasey et al., 2014). We found that the enrichment of P4RNA 3'-misincorporation at CG dinucleotides was eliminated in *ddm1 dcl3* compared to the *dcl3* single mutant (Figure 6B). We also examined trinucleotide enrichments in *ddm1 dcl3* at a set of CHG DMRs (Figure S7E), and observed a significant reduction at CHG but not at AHG, THG, or GHG (Figure S7F). Moreover, P4RNAs in *ddm1/dcl3* were slightly longer than those in *dcl3* single mutant (Figure S7G). Take together, these results suggest that DNA methylation itself is contributing to the pattern of 3'-misincorporation.

If P4RNAs are the precursors of siRNAs and DCL3 can cleave to some extent from the 3'-end of P4RNA, siRNAs should also contain some level of misincorporated bases at their 3' ends. By allowing for mismatches during genome mapping, we indeed found that siRNAs contain non-templated nucleotides that were enriched at the 3' end (Figure S6D). However, the proportion of siRNAs with 3'-mismatches (~1%) was far lower than that of

P4RNAs (~50%, Figure 5B), suggesting that far fewer siRNAs are processed from the 3' ends of P4RNAs than from the 5' ends, possibly due to the preference of 5'-adenine by DCL3 (Nagano et al., 2014). We also observed enrichment of CG sites at the 3' end of siRNAs that showed misincorporation (Figure S6E), again consistent with some level of processing of the 3' end of P4RNAs into siRNAs.

Since DCL3 appears to process from both ends of the double-stranded Pol IV/RDR2-derived RNA, and since P4RNAs are enriched for adenines at their 5' ends and misincorporated nucleotides at their 3' ends (Figure 7A), two predictions are that siRNAs with 5' adenines should have lower than average 3' misincorporation, and siRNAs with 3' misincorporation should have lower than average 5' adenine content. Indeed, we found that siRNAs with 5' adenines had 50% lower misincorporation than siRNAs with other 5' nucleotides (Figure 7B), and siRNAs with misincorporated nucleotides showed a lower 5' adenine content than those with perfect matches to the genome (Figure 7C). These results further support that P4RNAs containing misincorporated 3' nucleotides are processed into siRNAs. In summary, our results support that three different mechanisms can contribute the formation of the 3' end of P4RNAs - termination or 3' end processing at sequences enriched for ACU at the last three positions of the P4RNA, termination associated with misincorporation at methylated cytosines, and termination associated with misincorporation at other nucleotides.

DISCUSSION

The results of this study support the general scheme that P4RNAs are first transcribed by Pol IV, made double stranded by RDR2, and diced by DCL3 and other dicers to make the siRNAs which are loaded into AGO4. DCL3 cleavage produces an siRNA duplex with symmetric structure (Matzke and Mosher, 2014), but only one strand (the guide strand) is retained in AGO4, while the other strand (the passenger strand) is cleaved by AGO4 and degraded (Ye et al., 2012). Little is known about guide-strand selection of 24-nt siRNAs, and it remains unclear why, at many loci, siRNAs predominately match to one strand, and direct strand-specific methylation (Zhong et al., 2014). Our analysis revealed that the strand biases for P4RNAs and siRNAs

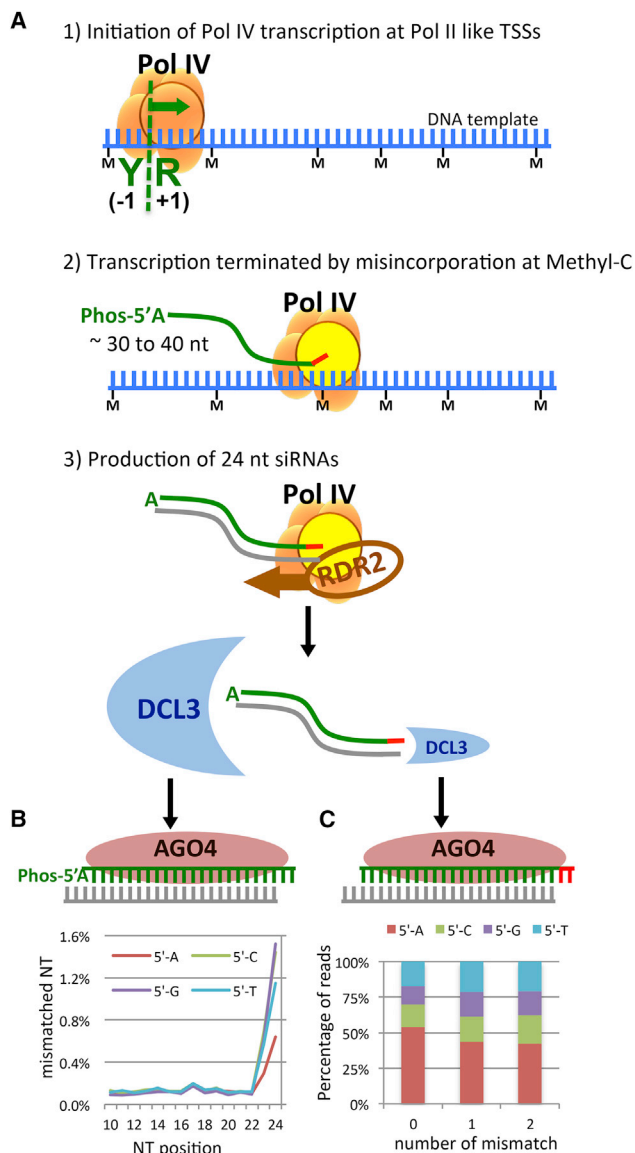


Figure 7. “One Precursor, One siRNA” Model for the Biogenesis of Pol IV Dependent 24-nt siRNAs

(A) Pol IV transcription is initiated at Pol II-like TSSs. A short RNA of ~30 to 40 nt (with 5'-adenine preference) is produced by Pol IV at heterochromatic regions; misincorporation then occurs at the cytosine position (red bar) and terminates Pol IV transcription. This process typically yields a P4RNA with 5' adenine and 3' misincorporation. P4RNA then goes through processing that involves RDR2 synthesizing the complementary strand (gray), DCL3 cutting from the 5' end (major) or 3' end (minor) of the P4RNA, and loading of the P4RNA-derived siRNA strand into AGO4 as the guide-strand. Eventually one P4RNA precursor gives rise to one siRNA that is derived from either its 5' or 3' end. (“M” underneath the DNA template indicates DNA methylation at the heterochromatic region.)

(B) When DCL3 cuts from the 5' end of P4RNAs, the resulting siRNAs are more likely to carry a 5' adenine and be perfectly matched. Indeed, a reduction of 3' mismatches is observed for siRNAs with a 5' adenine.

(C) When DCL3 cuts from the 3' end of P4RNAs, the resulting siRNAs are more likely to carry the 3' misincorporation and less likely the 5' adenine. This is consistent with the observation that siRNAs with mismatch have lower percentage of 5' adenine.

See also Figure S5.

are highly positively correlated (Figures 1D and S1D), suggesting that the siRNAs derived from the Pol IV strand rather than the RDR2-synthesized complementary strand, are favored as the guide-strand siRNA. This bias appears to be partially accomplished by a strong bias for adenine to be present as the first nucleotide of P4RNAs, by the preference of DCL3 for 5' adenines for siRNA processing, and by the preference of AGO4 for loading siRNAs with a 5' adenine (Figure 7A). P4RNAs also appear to be processed to some extent at their 3' ends to yield siRNAs that have similar 3' end signatures as are present in P4RNAs, including either an enrichment for ACU sequences, or misincorporated nucleotides. In addition, Pol IV transcripts are so short that on average only one siRNA will arise from each P4RNA. Finally, the short nature of Pol IV transcripts, coupled with the preference of DCL3 for short dsRNAs, may serve to channel Pol IV/RDR2 products into DCL3 rather than other dicers, such as DCL4 that prefers long dsRNAs (Nagano et al., 2014). These results imply a “one precursor, one siRNA” model for processing of siRNAs from Pol IV to RDR2 to DCL3 to AGO4 (Figure 7A). Based on the patterns of TSSs, strand-bias and 3'-misincorporation we conclude that the majority of PATH reads at siRNA loci are P4RNA, but it is still possible that some of these PATH reads are RDR2-transcribed as they are genetically dependent on both RDR2 and Pol IV.

Because Pol IV transcripts are so short, the sites of Pol IV transcriptional initiation and termination are largely determining the positions of siRNAs in the genome. Production of short transcripts by Pol IV could provide for siRNA biogenesis while avoiding the risk of transcribing full-length transposable elements. In addition, the short length of Pol IV transcripts may help to prevent spreading of RdDM to flanking regions, and allow specific silencing of transposons that are close to genes.

Our proposed model for biogenesis of 24-nt siRNAs in *Arabidopsis* share characteristics with that of 21U-RNAs (or piRNAs) in *C. elegans*. 21U-RNAs are 21-nt in length and begin with a 5' uridine, and they are autonomously expressed from thousands of loci dispersed in two broad regions of chromosome IV (Batista et al., 2008; Ruby et al., 2006). Like P4RNAs, precursors of 21U-RNAs, which are ~26 nucleotide Pol II transcripts, are terminated by an unknown mechanism to produce unusually short transcripts (Gu et al., 2012). However, unlike P4RNAs, 21U-RNA precursors are processed at their 5' ends to remove two nucleotides, and transcription units usually contain a conserved motif 42-nt upstream of the mature piRNA (Batista et al., 2008; Gu et al., 2012; Ruby et al., 2006). In contrast we did not find evidence for sequence conservation upstream of P4RNA start sites (Figure S4D), and instead Pol IV appears to be recruited by epigenetic signals such as the binding of H3K9 methylation through the Pol IV interacting protein SHH1 (Law et al., 2013). Furthermore, while there is clear evidence for transposon-derived secondary siRNA production through RNA-dependent RNA Polymerase (RdRP) activity in *C. elegans* (22G-RNA) (Lee et al., 2012), as well as Zucchini-dependent, secondary phased piRNAs in mammals (Han et al., 2015; Mohn et al., 2015), there is little evidence to support the possibility of secondary siRNA production in the RdDM pathway.

The 5' ends of P4RNAs are very similar in sequence composition to that of Pol II transcripts, showing strong enrichment for

Y/R dinucleotides at the $-1/+1$ positions. However, as opposed to Pol II transcripts that begin with trimethylguanosine caps, the majority of P4RNAs have 5' monophosphates. Since RNA polymerases normally start transcribing with a triphosphate containing nucleotide, it is unclear how Pol IV transcripts acquire a 5' monophosphate. It is possible that Pol IV is able to initiate transcription with a nucleoside monophosphate as previously reported for some RNA polymerase (Martin and Coleman, 1989; Ranjith-Kumar et al., 2002), or the P4RNAs may be initially capped after which a de-capping enzyme converts the trimethylguanosine cap to a 5' monophosphate, or an unknown polyphosphatase-like enzyme may directly convert the 5'-triphosphate to a 5'-monophosphate. Additionally, since 5' monophosphate containing RNAs are often a target of 5' to 3' exonucleases, there may be mechanisms to protect P4RNA 5' ends until subsequent processing steps are completed. It is also unclear why P4RNA evolved to have 5' monophosphates rather than cap structures, but one possibility is that this helps the cell avoid inadvertently mistaking a P4RNA for a Pol II transcript in order to avoid translation of transposon RNAs.

Our observations suggest that a component of the mechanism by which Pol IV transcription terminates is that DNA methylation on the template strand causes misincorporation of inappropriate bases. The mechanism by which this happens is unclear, but it is known that DNA methylation can cause transcriptional elongation defects in *Neurospora* (Rountree and Selker, 1997). The preferential termination of Pol IV transcription near DNA methylation would promote siRNA generation near sites of preexisting DNA methylation, thereby creating a self-reinforcing loop in which siRNAs direct DNA methylation targeting and DNA methylation helps direct the location of further siRNA production.

EXPERIMENTAL PROCEDURES

Biological Materials

The mutant alleles of *nripd1-4*, *nripb2-3*, *rdr2-1*, *nrip(d/e)2*, *rdr2/dcl3*, and combinations of *dcl2/3/4* used in this study were in the background of *Arabidopsis thaliana* ecotype Columbia-0 and have been previously described (Henderson et al., 2006; Li et al., 2008; Pontier et al., 2005; Xie et al., 2004; Zheng et al., 2009). The *dcl3/nripd1* double mutant was obtained by crossing *dcl3-1* with *nripd1-4*. Plants were grown in a growth chamber with 16 hr of light or greenhouse condition for five weeks. Immature inflorescence tissues including inflorescence meristem and early stages floral buds (up to stage 11/12) were collected.

Sequencing of Small RNA, PATH mRNA, and BS-Seq Libraries

Total RNA was first treated with RiboMinus Plant Kit for RNA-Seq (Invitrogen A10838-08) to remove rRNA, followed by size selection of RNA on a 15% UREA TBE Polyacrylamide gel (Invitrogen, EC6885BOX). Gels containing RNA with size between 15- to 27-nt were kept for small RNA library, while gels containing 28- to ~300-nt RNA were kept for PATH library. After gel elution, library construction for both sRNA and PATH was done using the Illumina TruSeq Small RNA Sample Preparation Kit (RS-200-0012), except that at the final size selection step, PCR products were separated on a 6% TBE Polyacrylamide gel (Invitrogen, EC6265BOX) and selected for the range from 120- to ~1,000-bp for PATH library. Gel-eluted PCR products with different TruSeq index sequences were pooled and sent for Illumina sequencing. The mRNA library was constructed using the Illumina TruSeq RNA Sample Preparation Kit (RS-122-2001) according to the standard manual. PATH libraries were sequenced using either paired-end mode with length of read1 being 120-bp and length of read2 being 30-bp (PE120+30), or single-end 100-bp (SE100); while sRNA and mRNA libraries were sequenced with

single-end 50-bp (SE50). For BS-seq, DNA was isolated using the DNeasy Plant Mini Kit (QIAGEN #69104) according to manufacturer instructions and quantified using the Qubit dsDNA High Sensitivity Kit (Life Technologies #Q32851). Libraries were constructed with 30ng DNA using the Ovation Ultralow Methyl-Seq Library Systems (NuGEN #0335). Bisulfite conversion was done using the EpiTect Bisulfite Kit (QIAGEN # 59104). BS-seq Libraries were sequenced at a length of 50 bp. All sequencing was carried out on Illumina HiSeq machines at the Broad Stem Cell Research Center (BSCRC) sequencing core at University of California, Los Angeles.

Real-Time RT-PCR and RNA Gel Blots

Real-time RT-PCR to detect P4RNAs in various genotypes was performed as the following: Total RNA were extracted from 100 mg of flowers using Trizol (Ambion) with an extra step of 24:1 Chloroform: isoamyl alcohol to remove remaining phenol prior to isopropanol precipitation, the resuspended RNAs were cleaned up using Quick-RNA miniprep (Zymo research, USA) for further purification and complete gDNA removal. Then 5ug of the purified RNA from each sample was used for RT reaction with SuperScript III first-strand kit (Invitrogen, USA). 1uL of the RT reaction was used for real-time PCR using iQ SYBR green supermix (Biorad, USA). We used terminator exonuclease (Epicenter, USA) for the removal of RNA with 5'-monophosphate: 5ug of total RNA were treated for an hour at 30 degree with Terminator, and the control, non-treated, RNA with the same buffer and 50% glycerol instead of terminator exonuclease enzyme. After Terminator treatment the RNA were used for RT reaction as described above. Primer information can be found in Table S3.

RNA gel blotting was performed as previously described (Henderson et al., 2006) with LNA probes for detecting transposon regions and regular RNA probes for detecting 5S locus. Probe information can be found in Table S3.

Pol II ChIP-Seq and AGO4-RIP

ChIP-seq was performed as described previously (Johnson et al., 2014). 5 μ g Pol II antibodies (Abcam #ab817) were used for each ChIP. Libraries for Pol II ChIP-seq were generated using the Ovation Ultralow DR Multiplex System (NuGen #0330) and sequenced at a length of 50 bp. ChIP-seq data were visualized using ngsplot (Shen et al., 2014).

AGO4-RIP was performed as previously described (Ji et al., 2011) with commercial AGO4 antibody (Agrisera #AS09 617). RNA isolated from RIP was used for library construction with Illumina TruSeq Small RNA Sample Preparation Kit (RS-122-2001) according to the standard manual.

Bioinformatic Analysis

Data handling

In general, qseq files received from the sequencing core were demultiplexed with an in-house Perl script and converted to fastq files for downstream analysis. For small RNA and PATH data, original reads were first trimmed using Cutadapt (v1.4), then mapped to the reference TAIR10 genome using Bowtie (Langmead et al., 2009) allowing only one unique hit (-m 1). We allowed zero mismatch for small RNA mapping (-v 0) and up to three mismatches for PATH mapping (-v 3). mRNA data were mapped using Tophat (Trapnell et al., 2009) allowing two mismatches and only one unique hit.

Analysis of siRNAs and P4RNAs

We used a list of Pol IV dependent siRNA loci that is previously described (Law et al., 2013) and manually inspected and filtered out a few loci that are tRNA related (listed in Table S2). In brief, these are 200 base pair bins where siRNAs were significantly reduced in a Pol IV mutant compared to two replicates of wild-type controls (FRD < 10^{-10}). More details can be found in the "Identification of siRNA clusters" section of Methods in the Law et al. paper (Law et al., 2013). Our definition of P4RNAs is any 27+ nt PATH reads derived from these ~7,000 previously defined Pol IV dependent 24-nt siRNA loci. For the abundance calculation, sRNA reads with the length between 18 and 26 nt, and PATH reads with the length greater or equal to 27 nt were included (Table S1). 5'-Adenine preference was analyzed by calculating the nucleotide composition at each position (counting from 5' end) for Pol IV siRNAs and P4RNAs. Percentage of 3'-non-templated nucleotides in P4RNAs was calculated at each position by analyzing the mapping results with a custom Perl script with the focus on the "MD:Z" column in sam format. The length of the 3'-misincorporation was done using the seed mapping option of bowtie to

map the first 24 nt of P4RNA perfectly and then allow up to ten mismatches for 3' portion. This analysis confirmed that majority of P4RNAs carry one or two 3'-misincorporated nucleotides. Therefore we chose to allow three mismatches for P4RNA mapping for all analysis.

Normalization of P4RNA and sRNA Abundance

Abundances of P4RNA in each library were normalized to the total number of sequenced reads, and abundances of siRNAs were normalized to the sum of all miRNAs in each library. For example, in Figure 2B, in Col PATH library we obtained 161,241,523 reads in total and of which 63,739 were classified as sasRNA (27+ nt), while in dcl2/3/4 PATH library we obtained 89,226,121 reads in total and of which 4,283,063 were classified as sasRNA. Therefore the percentage of P4RNA in dcl2/3/4 compared to Col is calculated as percentage = $(4,283,063/89,226,121) / (63,739/161,241,523) = 121.43 \approx 120$ -fold.

BS-Seq Analysis and DMR Calling

We used public data of floral ddm1 and wild-type BS-seq (GSE52346), ddm1/dcl3 sRNA-seq (GSE57191), and dcl3 sRNA-seq (GSE62801). Analysis of the floral BS-seq libraries of wild-type and ddm1 mutant (Creasey et al., 2014) were performed using BSMAP (Xi and Li, 2009), allowing only uniquely mapped reads and discarded sibling PCR products, with the tolerance of 2 mismatches per 50 bp. DMR calling was performed as previously described (Stroud et al., 2013), with a more stringent criteria: 1) sum of all sequenced cytosines in the 100 bp bin need to be at least 100; 2) the difference in CG methylation at each bin needs to be at least 0.5. This filtering allows us to focus on regions that are not only highly methylated in wild-type but also lose methylation dramatically in the ddm1 mutant. CHG DMRs were filtered with similar criteria, with the minimal loss of CHG in ddm1 compared to wild-type being 0.4, and the count of covered cytosines in each bin no less than 50.

Nucleotide Composition and Enrichment

P4RNAs from a defined set of loci were measured for their 3'-end sequence composition. For those P4RNAs that contain the 3' misincorporation, different compositions of the reference sequence at the first mismatched nucleotide (mono-nucleotide), the last matched plus first mismatched (di-nucleotide), or the last two matched plus first mismatched (tri-nucleotide) were counted as the observed value. The mono-, di- or tri-nucleotide compositions of the DNA sequence (or K-mer) were calculated using Jellyfish (Marçais and Kingsford, 2011) and considered as expected value assuming P4RNAs end equally likely with any nucleotide. The observed value and expected value were both normalized to their own population, and the ratio of observed/expected is used to show the relative enrichment of certain mono-, di- or tri-nucleotide composition.

ACCESSION NUMBERS

The accession number for the sequencing data reported in this paper is GEO: GSE61439.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.032>.

AUTHOR CONTRIBUTIONS

J.Z., S.B., S.F., I.A., I.R.H., X.C., and C.T. performed experiments. T.-f.L., S.F., and J.G-B. provided materials. S.Y.P., L.L., and X.C. participated in the genetic experiments. S.E.J. and B.C.M. oversaw the study. J.Z., H.W., and W.L. analyzed data. J.Z. and S.E.J. designed the study and wrote the manuscript.

ACKNOWLEDGMENTS

We thank members of the Jacobsen lab for insightful discussions. We thank Molly Megraw for suggestions on TSS analysis and Mahnaz Akhavan for technical assistance. We thank Dr. Jeffrey A. Long and Jared Sewell for their advice and guidance in performing ChIP-seq experiments and preparing genomic libraries. J.Z. thanks Dr. Chomdao Chommy and Seksan Sapsubsakul for

encouragement. High-throughput sequencing was performed at the UCLA BSCRC BioSequencing Core Facility. This work was supported by NIH grant GM60398 to S.E.J. and NSF award 1051576 to B.C.M. S.B. is supported by a postdoctoral fellowship of the Swiss National Science Foundation. J.G-B is a Human Frontiers Science Program fellow (LT000425/2012-L). J.Z. is a Life Science Research Foundation postdoctoral fellow, sponsored by the Gordon and Betty Moore Foundation. S.E.J. is an Investigator of the Howard Hughes Medical Institute.

Received: June 5, 2015

Revised: August 14, 2015

Accepted: September 11, 2015

Published: October 8, 2015

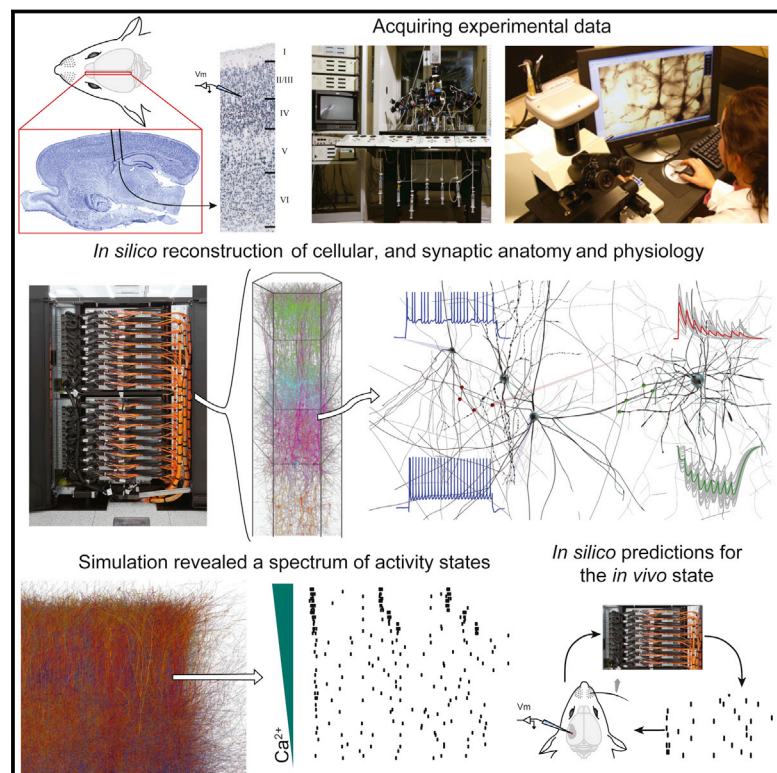
REFERENCES

- Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* 31, 67–78.
- Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452, 215–219.
- Colomé-Tatché, M., Cortijo, S., Wardenaar, R., Morgado, L., Lahouze, B., Sarazin, A., Etcheverry, M., Martin, A., Feng, S., Duvernois-Berthet, E., et al. (2012). Features of the Arabidopsis recombination landscape resulting from the combined loss of sequence variation and DNA methylation. *Proc. Natl. Acad. Sci. USA* 109, 16240–16245.
- Creasey, K.M., Zhai, J., Borges, F., Van Ex, F., Regulski, M., Meyers, B.C., and Martienssen, R.A. (2014). miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. *Nature* 508, 411–415.
- Cumbie, J.S., Ivanchenko, M.G., and Megraw, M. (2015). NanoCAGE-XL and CapFilter: an approach to genome wide identification of high confidence transcription start sites. *BMC Genomics* 16, 597.
- Gu, W., Lee, H.C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., Jr., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell* 151, 1488–1500.
- Haag, J.R., Ream, T.S., Marasco, M., Nicora, C.D., Norbeck, A.D., Pasa-Tolic, L., and Pikaard, C.S. (2012). In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol. Cell* 48, 811–818.
- Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* 348, 817–821.
- Havecker, E.R., Wallbridge, L.M., Hardcastle, T.J., Bush, M.S., Kelly, K.A., Dunn, R.M., Schwach, F., Doonan, J.H., and Baulcombe, D.C. (2010). The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22, 321–334.
- Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat. Genet.* 38, 721–725.
- Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science* 308, 118–120.
- Ji, L., Liu, X., Yan, J., Wang, W., Yumul, R.E., Kim, Y.J., Dinh, T.T., Liu, J., Cui, X., Zheng, B., et al. (2011). ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in Arabidopsis. *PLoS Genet.* 7, e1001358.
- Johnson, L.M., Du, J., Hale, C.J., Bischof, S., Feng, S., Chodavarapu, R.K., Zhong, X., Marson, G., Pellegrini, M., Segal, D.J., et al. (2014). SRA- and

- SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* 507, 124–128.
- Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J. (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* 37, 761–765.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature rev.* 11, 204–220.
- Law, J.A., Vashisht, A.A., Wohlschlegel, J.A., and Jacobsen, S.E. (2011). SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet.* 7, e1002195.
- Law, J.A., Du, J., Hale, C.J., Feng, S., Krajewski, K., Palanca, A.M., Strahl, B.D., Patel, D.J., and Jacobsen, S.E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* 498, 385–389.
- Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D., Jr., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* 150, 78–87.
- Li, C.F., Henderson, I.R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S.E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. *PLoS Genet.* 4, e27.
- Li, S., Vandivier, L.E., Tu, B., Gao, L., Won, S.Y., Li, S., Zheng, B., Gregory, B.D., and Chen, X. (2015). Detection of Pol IV/RDR2-dependent transcripts at the genomic scale in *Arabidopsis* reveals features and regulation of siRNA biogenesis. *Genome Res.* 25, 235–245.
- Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133, 523–536.
- Marçais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27, 764–770.
- Martin, C.T., and Coleman, J.E. (1989). T7 RNA polymerase does not interact with the 5'-phosphate of the initiating nucleotide. *Biochemistry* 28, 2760–2762.
- Matzke, M.A., and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature reviews* 15, 394–408.
- McCue, A.D., Panda, K., Nuthikattu, S., Choudury, S.G., Thomas, E.N., and Slotkin, R.K. (2015). ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *EMBO J.* 34, 20–35.
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., et al. (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell* 133, 116–127.
- Mohn, F., Handler, D., and Brennecke, J. (2015). Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* 348, 812–817.
- Nagano, H., Fukudome, A., Hiraguri, A., Moriyama, H., and Fukuhara, T. (2014). Distinct substrate specificities of *Arabidopsis* DCL3 and DCL4. *Nucleic Acids Res.* 42, 1845–1856.
- Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* 327, 335–338.
- Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C.S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 120, 613–622.
- Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* 19, 2030–2040.
- Ranjith-Kumar, C.T., Gutshall, L., Kim, M.J., Sarisky, R.T., and Kao, C.C. (2002). Requirements for de novo initiation of RNA synthesis by recombinant flaviviral RNA-dependent RNA polymerases. *J. Virol.* 76, 12526–12536.
- Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolić, L., and Pikaard, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* 33, 192–203.
- Rountree, M.R., and Selker, E.U. (1997). DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. *Genes Dev.* 11, 2383–2395.
- Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 127, 1193–1207.
- Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* 15, 284.
- Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V., and Jacobsen, S.E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* 152, 352–364.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.
- Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPPING program. *BMC Bioinformatics* 10, 232.
- Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* 2, E104.
- Yamamoto, Y.Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., and Abe, T. (2007). Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* 8, 67.
- Ye, R., Wang, W., Iki, T., Liu, C., Wu, Y., Ishikawa, M., Zhou, X., and Qi, Y. (2012). Cytoplasmic assembly and selective nuclear import of *Arabidopsis* Argonaute4/siRNA complexes. *Mol. Cell* 46, 859–870.
- Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.Y., and Chen, X. (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* 23, 2850–2860.
- Zhong, X., Du, J., Hale, C.J., Gallego-Bartolome, J., Feng, S., Vashisht, A.A., Chory, J., Wohlschlegel, J.A., Patel, D.J., and Jacobsen, S.E. (2014). Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* 157, 1050–1060.

Reconstruction and Simulation of Neocortical Microcircuitry

Graphical Abstract



Authors

Henry Markram, Eilif Muller, Srikanth Ramaswamy, Michael W. Reimann, ..., Javier DeFelipe, Sean L. Hill, Idan Segev, Felix Schürmann

Correspondence

henry.markram@epfl.ch

In Brief

A digital reconstruction and simulation of the anatomy and physiology of neocortical microcircuitry reproduces an array of in vitro and in vivo experiments without parameter tuning and suggests that cellular and synaptic mechanisms can dynamically reconfigure the state of the network to support diverse information processing strategies.

Highlights

- The Blue Brain Project digitally reconstructs and simulates a part of neocortex
- Interdependencies allow dense in silico reconstruction from sparse experimental data
- Simulations reproduce in vitro and in vivo experiments without parameter tuning
- The neocortex reconfigures to support diverse information processing strategies



Markram et al., 2015, Cell 163, 456–492
 October 8, 2015 ©2015 Elsevier Inc.
<http://dx.doi.org/10.1016/j.cell.2015.09.029>

Reconstruction and Simulation of Neocortical Microcircuitry

Henry Markram,^{1,2,19,*} Eilif Muller,^{1,19} Srikanth Ramaswamy,^{1,19} Michael W. Reimann,^{1,19} Marwan Abdellah,¹ Carlos Aguado Sanchez,¹ Anastasia Ailamaki,¹⁶ Lidia Alonso-Nanclares,^{6,7} Nicolas Antille,¹ Selim Arsever,¹ Guy Antoine Atenekeg Kahou,¹ Thomas K. Berger,² Ahmet Bilgili,¹ Nenad Buncic,¹ Athanasia Chalimourda,¹ Giuseppe Chindemi,¹ Jean-Denis Courcol,¹ Fabien Delalandre,¹ Vincent Delattre,² Shaul Druckmann,^{4,5} Raphael Dumusc,¹ James Dynes,¹ Stefan Eilemann,¹ Eyal Gal,⁴ Michael Emiel Gevaert,¹ Jean-Pierre Ghobril,² Albert Gidon,³ Joe W. Graham,¹ Anirudh Gupta,² Valentin Haenel,¹ Etay Hay,^{3,4} Thomas Heinis,^{1,16,17} Juan B. Hernando,⁸ Michael Hines,¹² Lida Kanari,¹ Daniel Keller,¹ John Kenyon,¹ Georges Khazen,¹ Yihwa Kim,¹ James G. King,¹ Zoltan Kisvarday,¹³ Pramod Kumbhar,¹ Sébastien Lasserre,^{1,15} Jean-Vincent Le Bé,² Bruno R.C. Magalhães,¹ Angel Merchán-Pérez,^{6,7} Julie Meystre,² Benjamin Roy Morrice,¹ Jeffrey Muller,¹ Alberto Muñoz-Céspedes,^{6,7} Shruti Muralidhar,² Keerthan Muthurasa,¹ Daniel Nachbaur,¹ Taylor H. Newton,¹ Max Nolte,¹ Aleksandr Ovcharenko,¹ Juan Palacios,¹ Luis Pastor,⁹ Rodrigo Perin,² Rajnish Ranjan,^{1,2} Imad Riachi,¹ José-Rodrigo Rodríguez,^{6,7} Juan Luis Riquelme,¹ Christian Rössert,¹ Konstantinos Sfyarakis,¹ Ying Shi,^{1,2} Julian C. Shillcock,¹ Gilad Silberberg,¹⁸ Ricardo Silva,¹ Farhan Tauheed,^{1,16} Martin Telefont,¹ Maria Toledo-Rodriguez,¹⁴ Thomas Tränkler,¹ Werner Van Geit,¹ Jafet Villafranca Díaz,¹ Richard Walker,¹ Yun Wang,^{10,11} Stefano M. Zaninetta,¹ Javier DeFelipe,^{6,7,20} Sean L. Hill,^{1,20} Idan Segev,^{3,4,20} and Felix Schürmann^{1,20}

¹Blue Brain Project, École polytechnique fédérale de Lausanne (EPFL) Biotech Campus, 1202 Geneva, Switzerland

²Laboratory of Neural Microcircuitry, Brain Mind Institute, EPFL, 1015 Lausanne, Switzerland

³Department of Neurobiology, Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

⁴The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

⁵Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

⁶Laboratorio Cajal de Circuitos Corticales, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, 28223 Madrid, Spain

⁷Instituto Cajal (CSIC) and CIBERNED, 28002 Madrid, Spain

⁸CeSViMa, Centro de Supercomputación y Visualización de Madrid, Universidad Politécnica de Madrid, 28223 Madrid, Spain

⁹Modeling and Virtual Reality Group, Universidad Rey Juan Carlos, 28933 Móstoles, Madrid, Spain

¹⁰Key Laboratory of Visual Science and National Ministry of Health, School of Optometry and Ophthalmology, Wenzhou Medical College, Wenzhou 325003, China

¹¹Caritas St. Elizabeth's Medical Center, Genesys Research Institute, Tufts University, Boston, MA 02111, USA

¹²Department of Neurobiology, Yale University, New Haven, CT 06510 USA

¹³MTA-Debreceni Egyetem, Neuroscience Research Group, 4032 Debrecen, Hungary

¹⁴School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, United Kingdom

¹⁵Laboratoire d'informatique et de visualisation, EPFL, 1015 Lausanne, Switzerland

¹⁶Data-Intensive Applications and Systems Lab, EPFL, 1015 Lausanne, Switzerland

¹⁷Imperial College London, London SW7 2AZ, UK

¹⁸Department of Neuroscience, Karolinska Institutet, Stockholm 17177, Sweden

¹⁹Co-first author

²⁰Co-senior author

*Correspondence: henry.markram@epfl.ch

<http://dx.doi.org/10.1016/j.cell.2015.09.029>

SUMMARY

We present a first-draft digital reconstruction of the microcircuitry of somatosensory cortex of juvenile rat. The reconstruction uses cellular and synaptic organizing principles to algorithmically reconstruct detailed anatomy and physiology from sparse experimental data. An objective anatomical method defines a neocortical volume of $0.29 \pm 0.01 \text{ mm}^3$ containing ~31,000 neurons, and patch-clamp studies identify 55 layer-specific morphological and 207 morpho-electrical neuron subtypes. When digitally reconstructed neurons are positioned in the volume and synapse formation is restricted to biological bouton densities and numbers of synapses per connection,

their overlapping arbors form ~8 million connections with ~37 million synapses. Simulations reproduce an array of in vitro and in vivo experiments without parameter tuning. Additionally, we find a spectrum of network states with a sharp transition from synchronous to asynchronous activity, modulated by physiological mechanisms. The spectrum of network states, dynamically reconfigured around this transition, supports diverse information processing strategies.

INTRODUCTION

Since Santiago Ramón y Cajal's seminal work on the neocortex (DeFelipe and Jones, 1988; Ramón y Cajal, 1909, 1911), a vast number of studies have attempted to unravel its multiple levels

of anatomical organization (types of neurons, synaptic connections, layering, afferent and efferent projections within and between neocortical regions, etc.) and functional properties (neuronal response characteristics, synaptic responses and plasticity, receptive fields, functional neocortical columns, emergent activity maps, interactions between neocortical regions, etc.). However, there are still large gaps in our knowledge, especially concerning the anatomical and physiological organization of the neocortex at the cellular and synaptic levels.

Specifically, while neurons have been classified in terms of their electrophysiological behaviors (Connors and Gutnick, 1990; Kasper et al., 1994; McCormick et al., 1985), expression of different calcium-binding proteins and neuropeptides (Celio, 1986; DeFelipe, 1993; Gonchar and Burkhalter, 1997; Kawaguchi and Kubota, 1997; Toledo-Rodriguez et al., 2005) and morphological features (Kisvárdy et al., 1985; Larkman, 1991a; Tamás et al., 1998; Wang et al., 2002), there is still no consensus on an objective and comprehensive classification of neuron types. Although the distribution of protein and genetic markers for different neurons (Grange et al., 2014; Hendry et al., 1989; Kawaguchi and Kubota, 1997; Meyer et al., 2002; Toledo-Rodriguez et al., 2004) and the relative proportions of some morphologically and electrically classified neurons (Beaulieu and Colonnier, 1983; Cauli et al., 1997; Hendry et al., 1984; Meyer et al., 2010a; Rudy et al., 2011) have been described, we lack a comprehensive view of the number of each type of neuron in each layer. Since the advent of paired recording techniques, several studies have characterized the anatomical and physiological properties of synaptic connections between some types of neurons (Cobb et al., 1997; Feldmeyer et al., 1999; Frick et al., 2008; Gupta et al., 2000; Mason et al., 1991; Reyes et al., 1998; Thomson et al., 1993), but a large proportion have yet to be studied. Although labeling with retrograde and anterograde tracers and trans-synaptic viral vectors, imaging with array tomography, and saturated reconstruction with electron microscopy have made it possible to begin mapping pre- and postsynaptic neurons for individual neocortical neurons (Boyd and Matsubara, 1991; Callaway, 2008; Glenn et al., 1982; Kasthuri et al., 2015; Killackey et al., 1983; Micheva and Smith, 2007; Micheva et al., 2010; Wickersham et al., 2007), we know neither the numbers and types of the pre- and postsynaptic neurons associated with any specific neuron type nor the numbers and locations of the synapses that they form with their immediate neighbors.

At a functional level, there have been many investigations of emergent behavior in neocortical slices (Cunningham et al., 2004; Mao et al., 2001; McCormick et al., 2003; Sanchez-Vives and McCormick, 2000; Yuste et al., 1997), correlated activity (Hasenstaub et al., 2005; Livingstone, 1996; Salinas and Sejnowski, 2001; Shu et al., 2003; Silberberg et al., 2004; Singer, 1993), and the functional impact of individual neurons across cortical layers (Sakata and Harris, 2009; Schroeder and Foxe, 2002; Silva et al., 1991; Steriade et al., 1993), as well as in vivo activity in somatosensory and other cortical areas (Chen et al., 2015; Klausberger et al., 2003; Leinekugel et al., 2002; Luczak et al., 2007; Reyes-Puerta et al., 2015; Wilson et al., 2012). However, we still lack an understanding of the cellular and synaptic mechanisms and the role of the different layers in the simplest

of behaviors, such as correlated and uncorrelated single-neuron activity and, more generally, synchronous and asynchronous population activity. For example, it is known that different types of neurons are connected through synapses with different dynamics and strengths, strategically positioned at different locations on the neurons' dendrites, somata, and axons, but the functional significance of this organization remains unclear. Computational approaches that abstract away this level of biological detail have not been able to explain the functional significance of such intricate cellular and synaptic organization. Although future experimental research will undoubtedly advance our knowledge, it is debatable whether experimental mapping alone can provide enough data to answer these questions.

Here, we present a complementary algorithmic approach that reconstructs neuronal microcircuitry across all layers using available sparse data and that leverages biological principles and interdependencies between datasets to predict missing biological data. As a test case, we digitally reconstructed a small volume of tissue from layers 1 to 6 of the hind-limb somatosensory cortex of 2-week-old Wistar (Han) rat. This model system was chosen not only because it is one of the most comprehensively characterized in the neocortex, but also because experimental data on its cellular and synaptic organization are readily available and validation experiments are relatively easy to perform. In brief, we recorded and digitally reconstructed neurons from in vitro brain slices and classified the neurons in terms of well-established morphological types (m-types; Figure 1A), positioned the neurons in a digital volume of objectively defined dimensions according to experimentally based estimates of their layer specific densities (Figure 1B), and reconstructed the connectivity between the neurons (Figure 1C). Neurons were then classified into electrical types (e-types), using an extended version of the classification proposed in the Petilla convention (Ascoli et al., 2008), and models were produced that captured the characteristic electrical behavior of each type (Figure 1D); similarly, synapses were modeled to capture the characteristic synaptic dynamics and kinetics of particular synapse types (s-types; Figure 1E). Finally, we constructed a virtual slice and reconstructed thalamic input using experimental data (Figure 1F; Meyer et al., 2010b).

This approach yielded a first-draft digital reconstruction of the microcircuitry, which was validated against a multitude of experimental datasets not used in the reconstruction. The results suggest that it is possible to obtain dense maps of neural microcircuitry without measuring every conceivable biological parameter and point to minimal datasets required, i.e., strategic data. Integrating complementary, albeit sparse, datasets also makes it possible to reconcile discrepancies in the literature, at least partially addressing the problem of data quality and reproducibility. Simulations exploring some of the emergent behaviors of the reconstructed microcircuitry reproduce a number of previous in vitro and in vivo findings and provide insights into the design and functioning of neocortical microcircuitry. The experimental data, the digital reconstruction, and the simulation results are available at the Neocortical Microcircuit Collaboration Portal (NMC Portal; <https://bbp.epfl.ch/nmc-portal>; see Ramaswamy et al., 2015).

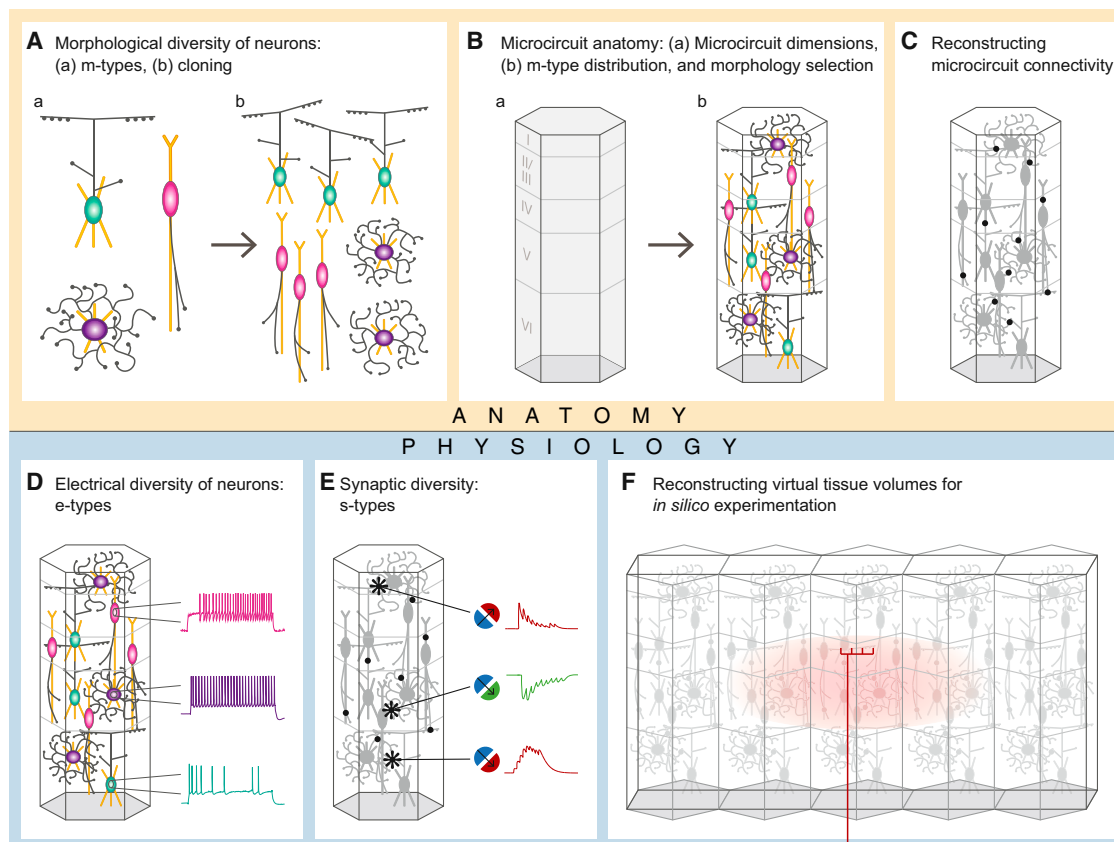


Figure 1. Workflow for Data-Driven Reconstruction of Neocortical Microcircuitry

(A) Morphological diversity of neurons. (a) Identify the morphological diversity in the neocortical microcircuit (m-types). (b) Repair and then clone the various m-types with statistical variations to enrich the number of exemplars.

(B) Microcircuit anatomy. (a) Define the spatial dimensions of a unitary microcircuit. (b) Assemble individual neurons in 3D space according to the frequency of occurrence of each m-type per layer, selecting the appropriate m-type instance that satisfies laminar constraints on the axonal and dendritic distribution.

(C) Reconstructing microcircuit connectivity. Derive the number and location of synaptic contacts formed between all neurons in the microcircuit, based on a series of synaptic connectivity rules.

(D) Electrical diversity of neurons. Map and model the electrical types (e-types) of each m-type to account for the observed diversity of morpho-electrical subtypes (me-types).

(E) Synaptic diversity of neurons. Map and model the diversity of synaptic types (s-types) observed between pre-post combinations of me-types, according to rules derived from synaptic physiology.

(F) Reconstructing virtual tissue volumes. Apply the above strategy to reconstruct defined circuit volumes (microcircuits, slices, mesocircuits) for *in silico* experiments; insert synapses formed by thalamocortical fibers for stimulation experiments.

RESULTS

Neuron-type Nomenclature

Neurons differ in terms of their location in the brain, morphology, electrical properties, projections, and the genes and proteins that they express (for reviews, see [Harris and Shepherd, 2015](#); [Markram et al., 2004](#)). The combination of these properties implies an immense diversity of neuron types. Given the lack of sufficient data for other dimensions, the neuronal classification used for this first-draft digital reconstruction considered only layer, local morphology, and electrophysiology. Naming of morphological types was based on the most common names used over the past century ([Connors and Gutnick, 1990](#); [DeFelipe, 1993](#); [DeFelipe et al., 2013](#); [Douglas and Martin, 2004](#);

[Fairén et al., 1984](#); [Hestrin and Armstrong, 1996](#); [Kawaguchi and Kubota, 1997](#); [Kisvárdy et al., 1985](#); [Oberlaender et al., 2012](#); [Somogyi et al., 1982, 1998](#); [Svoboda et al., 1997](#); [Szabadics et al., 2006](#)), extended with a layer prefix (e.g., Layer_Morphology, L5_MC for layer 5 Martinotti cells). Electrical types, based on the Petilla convention ([Ascoli et al., 2008](#)), were treated as subtypes, (e.g., L5_MC_NAC for the non-accommodating subtype; see [Experimental Procedures](#)). When whole-brain axonal tracing data for a sufficient number of projecting neurons becomes available (e.g., L5_TTPC_CP and L5_TTPC_CT to represent cortico-pontine and cortico-tectal subtypes; [Hallman et al., 1988](#); [Wang and McCormick, 1993](#); for a review, see [Ramswamy and Markram, 2015](#)), the proposed classification can be extended to include projection subtypes. Similarly, when there

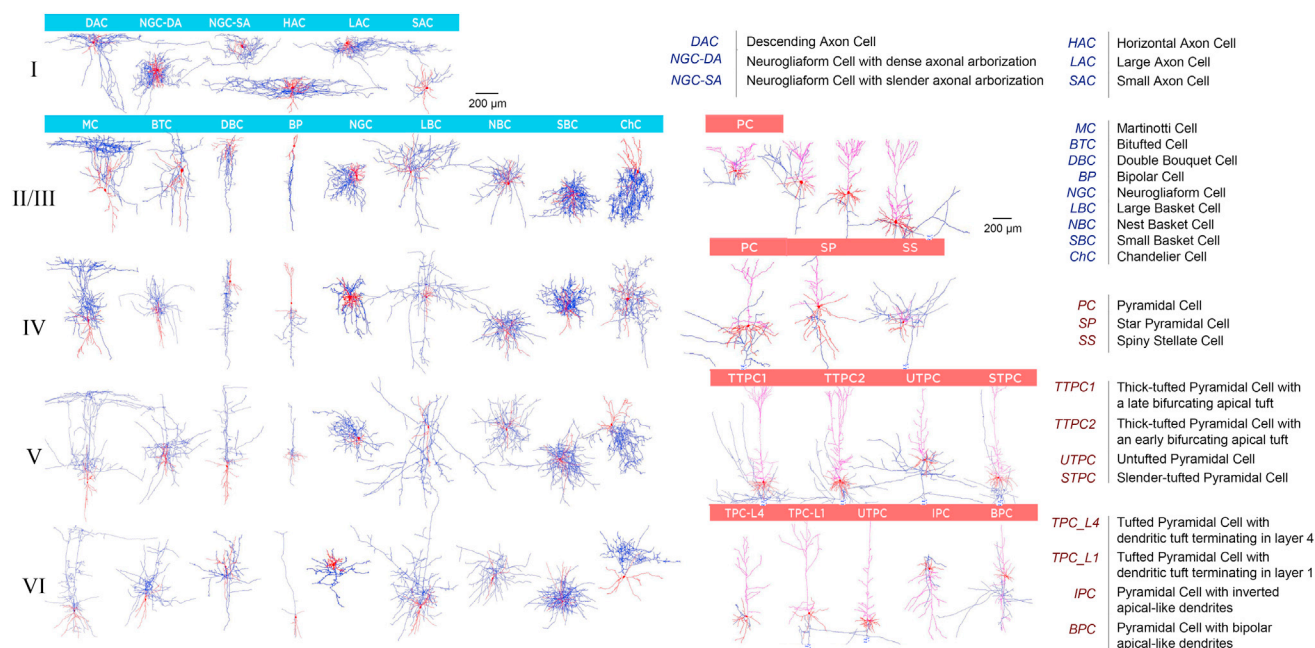


Figure 2. Table of Neocortical Neuronal Morphologies

Exemplar 3D reconstructions of 55 m-types. Morphologies in L2 and L3 are not separated. Axon in blue, dendrites in red. Full morphologies are not always shown. See also [Figure S1](#) for average arbor densities of each m-type and [Figure S2](#) for objective classification of m-types and details of the morphology cloning process. See also [Movie S1A](#).

are sufficient single-cell gene and protein expression data to systematically identify cells, it can be extended to include molecular subtypes. The abbreviations used for each m-type are provided in [Figure 2](#). A mapping between the nomenclature used in this study and alternative names present in the literature is provided in [Table 1](#).

Morphological Diversity of Neocortical Neurons

We recorded and labeled >14,000 neurons from all six layers in the somatosensory cortex of P14 male Wistar (Han) rats, using patch-clamp electrodes in *in vitro* slices. Of these neurons, 2,052 were sufficiently well stained to allow expert classification into m-types, based on well-established characteristic features of their dendritic and axonal arbors, a procedure initiated by early neuroanatomists and still in use today ([Fairén et al., 1984](#); [Karagiannis et al., 2009](#); [Karube et al., 2004](#); [Kawaguchi and Kubota, 1997](#); [Kisvárdy et al., 1985](#); [Larkman, 1991a](#); [Perrenoud et al., 2013](#); [Peters and Kaiserman-Abramof, 1970](#); [Ramón y Cajal, 1909, 1911](#); [Somogyi et al., 1982, 1998](#); [Wang et al., 2004](#); [Yuste, 2005](#)). We were able to digitally reconstruct a subset of 1,009 of these neurons. This allowed validation of the expert classification using an objective method (see below) based on clustering of characteristic features and provided the initial pool of digital neuron models needed to reconstruct the microcircuitry. In a few cases, we had no morphological reconstructions for rare m-types known to be present in the microcircuitry (L5_BP, L5_ChC, L6_NGC; [Oláh et al., 2007](#); [Szabadics et al., 2006](#)). These were represented using exemplars of the same morphology from neighboring layers. Although L6 horizontal and

sub-plate pyramidal cells (L6_HPC and L6_SPC) were present in the dataset and have also been reported in the literature ([Ghosh and Shatz, 1993](#); [Hevner et al., 2001](#)), the quality of the stains was not sufficient for reliable reconstruction. These morphologies are not represented in the first draft.

Aggregating morphological reconstructions and reports in the literature, we distinguished 55 m-types (65 if layers 2/3 are considered separately and 67 if L6_HPC and L6_SPC are also considered; [Figure 2](#)). Inhibitory types are mostly distinguished by axonal features and excitatory types by dendritic features (for reviews, see [Markram et al., 2004](#); [Ramaswamy and Markram, 2015](#); [Spruston, 2008](#)). [Figure S1](#) shows overlays of multiple exemplars of each of the 55 major m-types, and [Figures S2A](#) and [S2B](#) illustrate the objective classification. While in some cases, it might have been possible to introduce a finer separation between m-types, this would have limited the size of the samples for individual types, reducing the reliability of the classification.

The same inhibitory types were present in all layers except layer 1, which contained a unique set of inhibitory neuron types. Pyramidal cell morphologies varied across layers ([Figure 2](#), right) and also with depth within layer, as illustrated by the diversity of L23 PCs ([Figure 2](#), upper-right). The number of pyramidal cell types, as defined by their local morphology, increased from upper to lower layers. Several types of interneurons (e.g., LBC and DBC) had axonal arbors that tended to descend to deeper layers when they were in upper layers and to ascend to upper layers when they were in deeper layers. Consistent with this trend, one type of pyramidal cell (L6_IPC) also had inverted axonal arbors.

Table 1. Relation of Interneuron Classes to Classification Schemes Found in the Literature

Morphological Type	Neurogliaform Cell (NGC)	Small Basket Cell (SBC)	Double Bouquet Cell (DBC)	Bipolar Cell (BP)	Martinotti Cell (MC)	Bitufted Cell (BTC)	Large Basket Cell (LBC)	Nest Basket Cell (NBC)	Chandelier Cell (ChC)
Other morphological classifications	Dwarf cell, button-type cell	Clutch cell	Bitufted cell/ interneuron, horse-tail cell	Bitufted cell/ interneuron	Bitufted cell/ interneuron	Bitufted interneuron	Common basket cell, typical basket cell	Willow cell, arcade cell, shaft-biased cell, atypical basket cell	Axo-axonic cell
Predominantly expressed Ca ²⁺ -binding proteins and peptides	CB (–), NPY (+), PV (–), VIP (–), CR (–) SOM (–)	CB (++) NPY (+), PV (–), VIP (+++), CR (–) SOM (++)	CB (+), NPY (–), PV (–), VIP (+++), CR (+) SOM (++)	CB (–), NPY (–), PV (–), VIP (+++), CR (++) SOM (++)	CB (++) NPY (+++), PV (–), VIP (–), CR (–) SOM (+++)	CB (++) NPY (+), PV (–), VIP (+), CR (++) SOM (++)	CB (++) NPY (+), PV (+++), VIP (+), CR (+) SOM (–)	CB (++) NPY (+), PV (+++), VIP (+), CR (++) SOM (–)	CB (+), NPY (–), PV (–), VIP (–), CR (–) SOM (–)
Electrical types	bNAC (7%), cNAC (79%), cSTUT (7%), cAC (7%)	bNAC (36%), cAC (36%), dNAC (29%)	bAC (9%), cAC (9%), bIR (37%), cIR (18%), bNAC (9%), cNAC (9%), bSTUT (9%)	bAC (7%), cAC (29%), bIR (14%), cIR (14%), bNAC (29%), cNAC (7%)	bAC (37%), cAC (3%), bIR (11%), cIR (3%), bSTUT (4%), cSTUT (37%)	bAC (17%), cAC (67%), cNAC (17%)	bAC (6%), cAC (12%), cNAC (17%), dNAC (17%), cIR (6%), dSTUT (24%)	bAC (6%), cIR (7%), bIR (6%), cNAC (20%), bSTUT (13%), cSTUT (20%), cAC (20%), dSTUT (7%)	cAC (38%), cNAC (38%), dNAC (25%)
Other electrical classifications	Non-fast spiking, late spiking	Fast spiking, non-accommodating, non-adapting	Irregular spiking, regular spiking non-pyramidal, adapting	Late spiking, regular spiking non-pyramidal, adapting	Regular spiking non-pyramidal, burst spiking non-pyramidal, low threshold spiking	Regular spiking non-pyramidal, adapting, burst spiking non-pyramidal	Fast spiking, non-accommodating, non-adapting	Fast spiking, non-accommodating, non-adapting	Fast spiking, late spiking, non-adapting

The terms used in this paper are in the first row, followed by other common names in the literature. Interneurons can be categorized according to which primary marker they express (calcium-binding proteins: parvalbumin [PV], calbindin [CB], and calretinin [CR]; neuropeptides: somatostatin [SOM], vasoactive intestinal polypeptide [VIP], neuropeptide Y [NPY], and cholecystokinin [CCK]). The mapping to serotonergic receptors (5HT_{3A}R) is not included since this was not assayed in the RT-PCR. We assign several possible electrical types to each morphological type, based on the Petilla convention, and show other names frequently used in the literature. See [Figure 4](#) for definitions of electrical types.

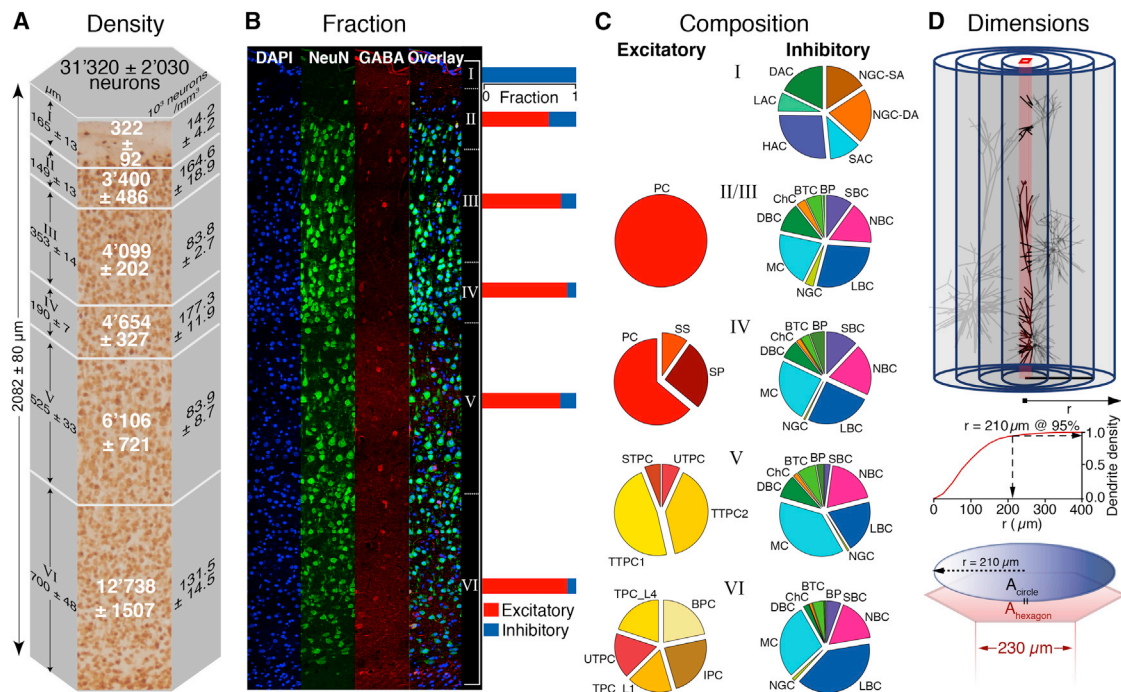


Figure 3. Neuron Densities and Composition and Microcircuit Dimensions.

(A) Neuron densities and numbers. Vertical thicknesses as determined by transitions in neuronal somata size and density in NeuN stained slices (six animals; mean \pm SD). Neuron densities and numbers (six animals; mean \pm SD).

(B) Neuron fractions. Confocal block imaging of dual immunohistochemical labeling. DAPI labels all cells (blue). NeuN labels all neurons (green), GABA labels all GABAergic cells including glia (red), dual GABA and NeuN labels only GABAergic neurons (green). Bars to the right show fractions of excitatory (red) and inhibitory (blue) neurons in each layer.

(C) m-type composition. Fractions of inhibitory (left) and excitatory (right) m-types per layer ($n = 2052$).

(D) Dimensions. The horizontal dimension was defined as the smallest circle required to attain maximal dendritic volume at a central minicolumn (brown, top); cut-off radius, 95% of the plateau volume ($r = 210 \mu\text{m}$, middle). To allow tiling, the circle was transformed into a hexagon, preserving the area. For m-type acronyms, see Figure 2.

See also Figure S3 for details on morphology placement and Figure S4 for validation of the composition. See also Movie S1B.

Using multiple exemplars obtained from different animals for each m-type, we developed a repair process to recover arbors cut during the slicing process, which was validated using *in vivo* reconstructed neurons (see [Experimental Procedures](#); [Anwar et al., 2009](#)). To generate an even larger pool of unique morphologies, we cloned multiple exemplars of each m-type (Figures S2C–S2F), jittering branch angles, and section lengths in the clones (see [Experimental Procedures](#)). The morphometric properties of the resulting population were validated against distributions of features obtained from reconstructed neurons (see [Experimental Procedures](#)). This approach allowed us to establish a dataset of neuronal morphologies (see [Movie S1A](#)) that respects biological variability. Software applications for repairing and cloning *in vitro* neuron morphologies and for automated classification of neurons into the 55 m-types are available through the NMC Portal.

Reconstructing Neuron Densities, Ratios, and Composition

Reconstruction began by specifying the dimensions of the microcircuit, the fractions of excitatory and inhibitory neurons, the proportions of each m-type, and the number of neurons of each

m-type. The height of the neocortex and heights of each layer were measured experimentally in six animals, yielding an average overall height of $2,082 \pm 80$ microns (mean \pm SD; $n = 6$; Figure 3A). Layer thicknesses were determined experimentally by measuring the location of transitions in cell densities and soma sizes in NeuN-stained tissue blocks (see [Experimental Procedures](#)). Fractions of excitatory and inhibitory neurons per layer (E-I fractions) were established by counting cells stained for DAPI (all cells), NeuN (all neurons), and GABA (all inhibitory neurons) in tissue blocks (Figure 3B; see [Experimental Procedures](#)). Overall, excitatory and inhibitory neurons represented $87\% \pm 1\%$ and $13\% \pm 1\%$ of the population, respectively, with a trend toward higher fractions of excitatory neurons in deeper layers (Figure 3B).

The m-type composition for all excitatory and all inhibitory neurons in each layer was obtained from the relative frequencies of each m-type in the experimental dataset of 2,052 classified neurons mentioned earlier (Figure 3C; see [Experimental Procedures](#)). It is not possible to exclude sampling bias in this dataset. However, since E-I fractions were obtained in an unbiased manner, any bias is restricted to the proportions of m-types within the excitatory and inhibitory neurons and does not affect the overall E-I balance.

The E-I fractions and m-type composition determined in this way are broadly consistent with previous reports (DeFelipe et al., 2002; Lefort et al., 2009). For example, it is well established that ~50% of inhibitory interneurons are basket cells (i.e., LBCs and NBCs—predominantly parvalbumin-positive cells; SBCs—predominantly vasoactive intestinal peptide (VIP)-positive cells; we found ~53%, see below), that Martinotti cells (i.e., predominantly somatostatin-positive cells; we found ~22%, see below) are frequent in all layers except L1, and that bitufted and bipolar cells (i.e., many of the calbindin and calretinin-positive cells) and double bouquet cells (i.e., many of the VIP-positive cells) are both found in layers 2–6. Other inhibitory interneuron types are also found in L2–L6 but less frequently (Kawaguchi and Kubota, 1997; Krimer et al., 2005; Meyer et al., 2011; Oláh et al., 2007; Sancesario et al., 1998; Somogyi et al., 1998; for a review, see Markram et al., 2004). Previously published neuron densities could not be used because they varied by a factor of two (40,000–80,000 neurons/mm³; Beaulieu, 1993; Cragg, 1967; DeFelipe et al., 2002; Keller and Carlson, 1999; Peters, 1987) and are too low to account for the number of synapses in the microcircuit (see below, “Digital Reconstruction of Connectivity”). We therefore performed new experiments, counting cells in NeuN-stained tissue blocks. The experiments yielded a mean cell density of $108,662 \pm 2,754$ neurons/mm³ (mean \pm SEM, $n = 6$; see Experimental Procedures), comparable to observations in rat barrel cortex (Meyer et al., 2010a). Neuron densities were highest in L4 (Figure 3A), consistent with previous studies (Meyer et al., 2010a).

Since hind-limb somatosensory cortex, unlike barrel cortex, has no anatomically defined horizontal columnar organization (Horton and Adams, 2005; Markram, 2008), we chose to define the radius of the microcircuit by placing reconstructed neurons in a cylindrical volume and determining the minimal radius where the density of dendrites saturates at the center (Figure 3D; 95% of the plateau value obtained at a radius of 210 μ m; see Experimental Procedures). We chose dendrites, as opposed to axons, because they only arborize locally. This convention, which yields a minimal radius that reflects saturated dendritic density along the central axis, could allow comparisons between microcircuits in different brain regions. It yields a radius similar to the horizontal extent of the dendrites of the largest neuron in the microcircuit (i.e., the L5_TTPC; for a review, see Ramaswamy and Markram, 2015) and is comparable with the dimensions of the barrels in the rodent barrel cortex (Meyer et al., 2010b; Wimmer et al., 2010). To allow tiling of multiple microcircuits while minimizing edge effects, the volume of the microcircuit was defined as a hexagonal prism (Figure 3D, bottom) with a cross-sectional area equal to that of the circle with the radius defined above and a height determined by the combined height of the layers.

With these densities, m-type composition, and circuit dimensions, we calculated the number of each m-type in each layer and in the whole microcircuit. To approximate inter-individual variation in layer dimensions and neuronal densities, we digitally reconstructed separate microcircuits corresponding to layer heights and densities measured in five animals (Bio1–Bio5). The five reconstructions had an average of $31,375 \pm 2,251$ neurons (mean \pm SD, $n = 5$), with the number of neurons increasing in each layer from L1 to L6. We then constructed an additional

microcircuit using the averaged data (BioM). To assess the variation introduced by the digital reconstruction process (stochastic variations in m-type composition, selection and positioning of model neurons, and synaptic connectivity [see below]), we reconstructed seven instances of each microcircuit (i.e., seven reconstructions each from Bio1–Bio5 and seven from BioM; 42 in total).

Positioning Morphologically Reconstructed Neurons

After establishing the dimensions of the microcircuit and the number of neurons belonging to each m-type in each layer, it was necessary to position each neuron in the digital reconstruction. Consistent with reports of weak minicolumnar organization in rodents, (Mountcastle, 1998), neurons were arranged in 310 minicolumns at horizontal positions drawn from 2D Gaussians around the center of each minicolumn, thus relaxing the strictness of the minicolumnar organization (see Experimental Procedures). The positions of the neurons along the vertical axis of the minicolumn were randomly chosen within each layer, using a space-filling algorithm to ensure that somata did not overlap (see Experimental Procedures).

Once the positions of the neurons were established, a second algorithm randomly selected a suitable morphology for each position from the top 8% of morphologies, scored by their match to typical patterns of arborization within and across layers (Figure S3; see Experimental Procedures). These patterns were manually annotated on each reconstructed neuron, based on the depth of the recorded neuron within each layer and cross-layer arborization patterns described in the literature (see Experimental Procedures and NMC Portal). Figure S4A illustrates the microcircuit at this stage of reconstruction (see also Movie S1B). The total lengths of axons and dendrites in the average microcircuit were 350 ± 4 m and 215 ± 3 m (mean \pm SD, $n = 7$), respectively.

Biological accuracy at this stage of the reconstruction was validated against two experimental datasets that had not been used thus far. The first tissue-level dataset provides in vitro immunohistochemical staining of 30 μ m sections for seven markers (calcium-binding proteins and neuropeptides) commonly used to label inhibitory interneurons (Figure S4B). The second cellular-level dataset provides estimated probabilities that the genes for these markers are expressed in specific m-types (Toledo-Rodriguez et al., 2005; Wang et al., 2002, 2004). We used the second dataset to add the markers to the model neurons. We then performed in silico immunohistochemical staining of the whole reconstructed tissue for each marker separately and compared the in silico stains against immunohistochemical stains from the first dataset. Although gene expression data are noisy and genes do not translate equally to protein levels, we found a reasonable correspondence between the numbers of neurons at different depths stained for specific markers in the in silico and the in vitro stains (regression, $r = 0.65$; Figure S4C). Furthermore, the layer-dependent pattern of in silico stained cells was consistent with previous staining experiments in this brain region (Ascoli et al., 2008; Condé et al., 1994; DeFelipe, 1993; Dumitriu et al., 2007; Gentet et al., 2010, 2012; Gonchar and Burkhalter, 1997; Gonchar et al., 2007; Kawaguchi and Kubota, 2002; Kawaguchi and Kubota, 1993,

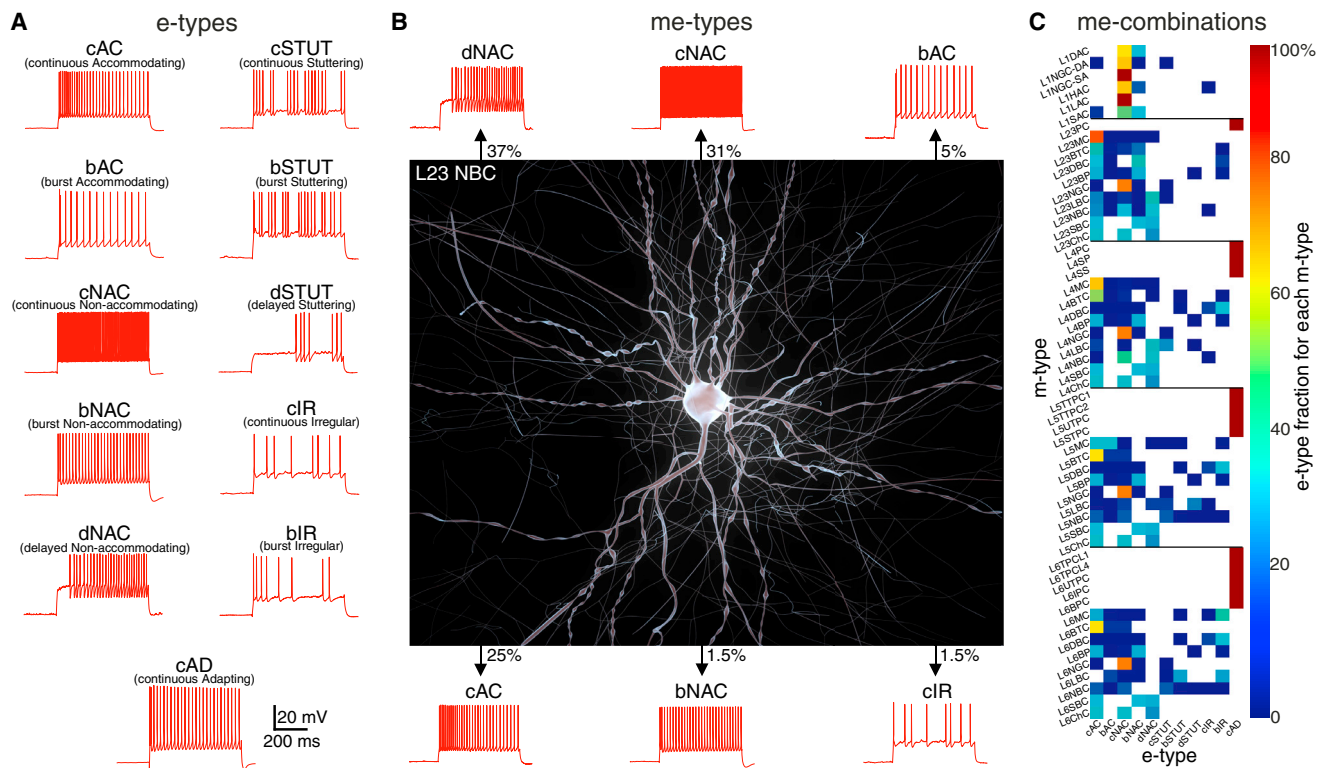


Figure 4. Table of Morpho-Electrical Neuron Types

(A) e-types. Diverse firing patterns in response to depolarizing step current injections in neocortical neurons. c, continuous; d, delayed; b, bursting. AC, accommodating; NAC, non-accommodating; STUT, stuttering; IR, irregular; AD, adapting.

(B) An exemplar neuron (L23NBC) with a diversity of e-types. Percentages indicate the relative frequency of e-type occurrence.

(C) Fractions of e-types (11 e-types) recorded experimentally in each of the 55 m-types, making up 207 me-types. Solid lines indicate layer boundaries. See Table 1 for relation of e-types to other classifications in the literature.

1997; McGarry et al., 2010; O'Connor et al., 2009; Packer and Yuste, 2011; Santana et al., 2013; see also NMC portal). The observed correspondence would be unlikely in the presence of major errors in neuron densities, m-type composition, or positioning of reconstructed neurons. However, the biological data are highly variable, and the validation of the inhibitory m-type composition used only a small proportion of markers reported in the literature. The reconstruction should thus be considered as a first draft, to be refined as it is challenged with additional markers.

Morpho-Electrical Composition

We applied a standardized battery of stimulation protocols (Le Bé et al., 2007; Wang et al., 2002, 2004) to >3,900 neurons from all layers, recording and analyzing their responses. The neurons were classified using quantified features of the neuronal response to step current pulses, according to the criteria established by the Petilla convention (Ascoli et al., 2008; Figure 4A, top), with the exception of stuttering cells, which were considered as a separate class (see Druckmann et al., 2013). Since no significant bursting behavior was observed in excitatory m-types from animals of the age used in this study, all excitatory m-types were classified as continuous adapting (cAD) neurons (Figure 4A, bottom). Using this feature-based

classification scheme, we identified 11 e-types (10 inhibitory e-types and 1 excitatory e-type) (Figure 4A; see Experimental Procedures). Objective clustering of the same features produced a similar classification, validating the original classification scheme (Druckmann et al., 2013). The fact that the e-types identified in this way have characteristic ion channel profiles provides further evidence for their distinctive identity (Khazen et al., 2012; Toledo-Rodriguez et al., 2004).

Most inhibitory m-types expressed multiple e-types (Figure 4B), consistent with previous observations (Ascoli et al., 2008; Cauli et al., 2000; Nelson, 2002; Toledo-Rodriguez et al., 2005). Combining m- and e-types yielded 207 morpho-electrical types (me-types), providing an integrated view of the morpho-electrical diversity of the microcircuit (Figure 4C). A dataset of 511 morphologically and electrically classified inhibitory neurons was used to determine the relative proportion of e-types for each inhibitory m-type (in a layer-dependent manner for m-types with sufficient samples and otherwise in a layer-independent manner; Figure 4C, color map; see Experimental Procedures). The relative proportions were combined with neuron densities to calculate the number of neurons for each me-type in each layer. The resulting diversity and spatial distribution of inhibitory e-types is illustrated in Figure 5A. This integrated view of the microcircuitry reveals that, at this age, the most common inhibitory

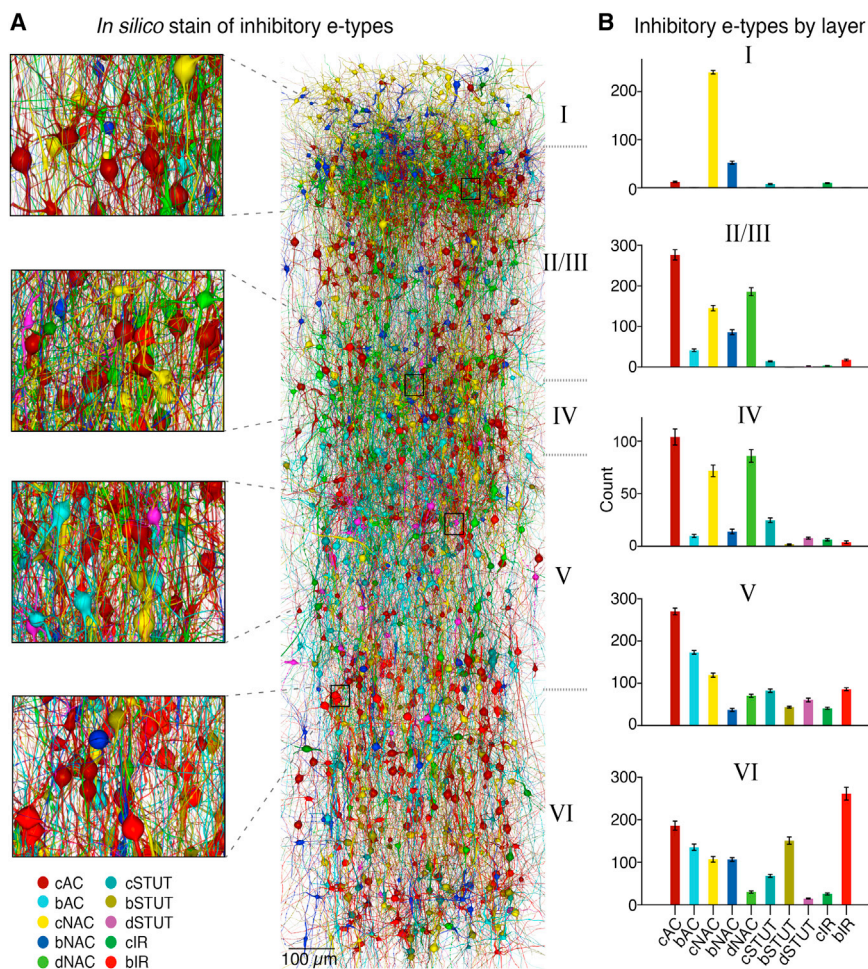


Figure 5. Layer-Dependent Distribution of Inhibitory e-Types

(A) *In silico* “brainbow” staining of a random selection of inhibitory morphologies, colored by e-type.

(B) Layer-wise distribution of inhibitory e-types ($n = 35$ reconstructions; mean \pm SD). See Figure 4 for definitions of e-types.

e-type is cAC, followed by cNAC and dNAC, and that stuttering and irregular e-types (cSTUT, bSTUT, dSTUT, cIR, and bIR) are relatively rare (Figure 5B). Inhibitory e-types with regular firing patterns (cAC, bAC, cNAC, bNAC, and dNAC) occur more frequently in superficial layers, whereas e-types with irregular firing patterns (cSTUT, bSTUT, dSTUT, cIR, bIR) are more common in deep layers (Figure 5B).

Digital Reconstruction of Connectivity

We developed an algorithmic approach to reconstruct synaptic connectivity between neurons in a companion study (Reimann et al., 2015). The approach is based on five rules of connectivity described in the Experimental Procedures and validated in Reimann et al. (2015). We implemented these rules in four stages that yield plausible multi-synapse connections, consistent with the rules and constrained by experimental bouton densities (Figure 6A).

The algorithm predicts the characteristics of multi-synapse connections between pairs of neurons that belong to specific m-types (Figure 6B). We have previously shown that these predictions faithfully reproduce detailed anatomical data on connectivity between L5 thick-tufted PCs (number of synapses and locations; Ramaswamy et al., 2012) and for a number of

other connection types (synapse locations; Hill et al., 2012). We now show that they reproduce the connectivity (numbers and locations) of all connection types that have been studied experimentally (see NMC Portal). For example, the anatomy of *in silico* synaptic connections between L5 Martinotti cells and L5 thick-tufted PCs (Figure S5) compares well with available experimental data (Silberberg and Markram, 2007). The algorithm provides detailed anatomical predictions for connection properties, which it has not yet been possible to measure experimentally (e.g., numbers of source and target cells and synapses) (Figure S5). The reconstruction also allows studies of neurons involved in polysynaptic pathways (see NMC Portal) forming known motifs (Honey et al., 2007; Perin et al., 2011; Silberberg, 2008; Sporns and Kötter, 2004).

The algorithm yields 1,941 biologically plausible multi-synapse connection types (out of a theoretical 3,025) that are consistent with the connectivity principles described above.

Figure 7 shows the predicted average number of synapses formed by each potentially viable connection type (Figure 7A) as well as their predicted average connection probabilities (Figure 7B). The predicted number of synapses/connection is 4.5 ± 0.1 (3.6 for excitatory connections, 13.9 for inhibitory connections; $n = 35$). We also predict 27,625 types of connection between neurons of different me-types (see NMC Portal).

On average, each neuron innervates 255 ± 13 other neurons belonging to $32\% \pm 1\%$ of m-types, forming an average of $1,145 \pm 75$ synapses per neuron present in the microcircuit (Figure S6A; mean \pm SD, across the 35 Bio1-5 reconstructions; all neurons sampled). As a population, the neurons belonging to a given m-type innervate $63\% \pm 6\%$ of the m-types in the microcircuit. The individual reconstructions (Bio1-5) yield an average of 638 ± 74 million appositions and 36.7 ± 4.2 million synapses (27.0 ± 2.9 million excitatory and 9.7 ± 1.5 million inhibitory). Taken together, the neurons of the microcircuit form 8.1 ± 0.9 million connections. Figure 7C and Table S1 provide a first view of the connectivity between neurons of the neocortical microcircuit. Analyzing these data, we find that, at this age, the fraction of excitatory synapses (red) increases from L1 to L6 (Figure S6B). At later ages, this trend may change as axons

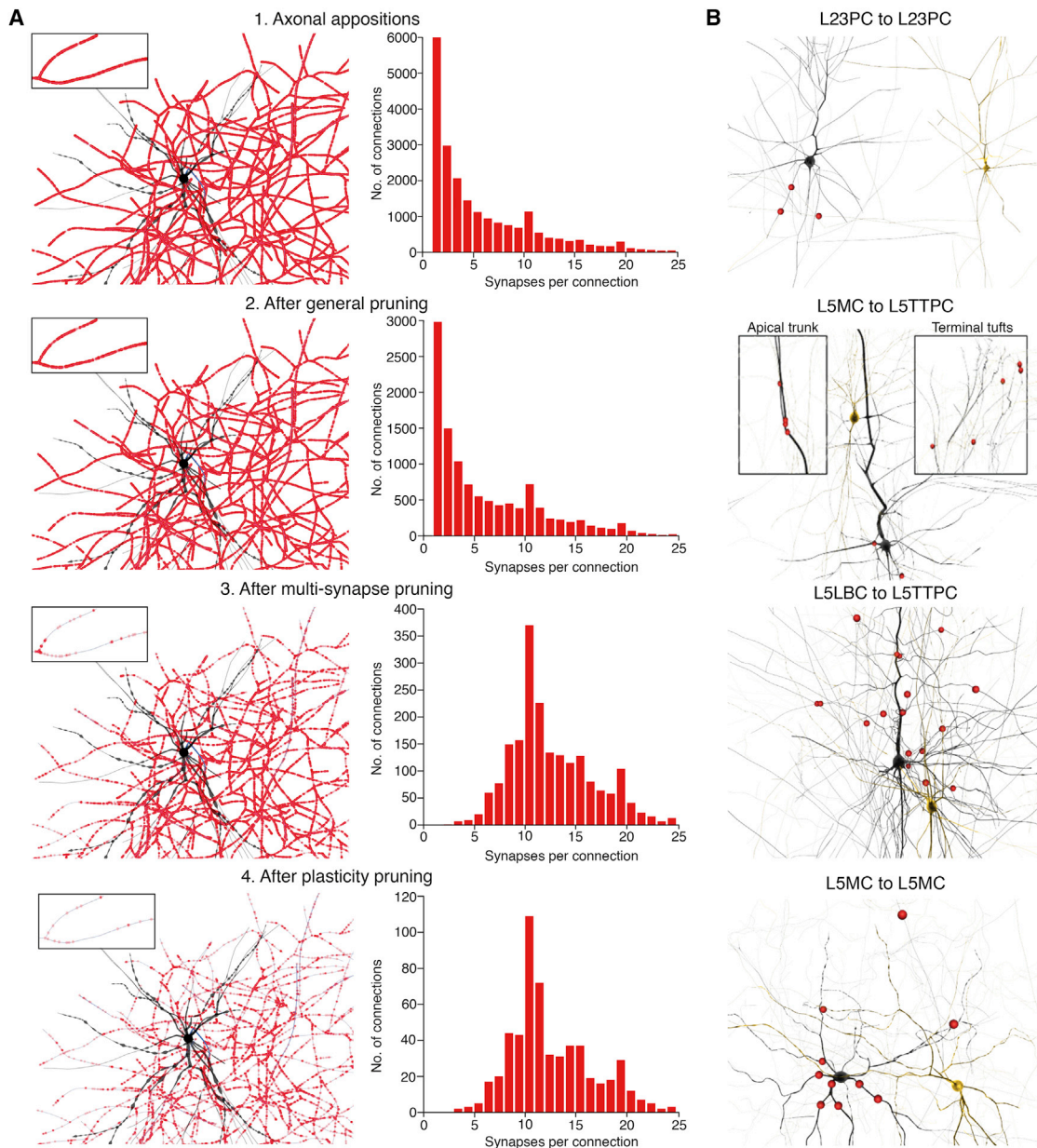


Figure 6. Reconstructing Connectivity

(A) Four-step algorithm to convert putative axo-dendritic appositions into functional synapses. (1) Axonal appositions. For an exemplar L23SBC (left, soma and dendrites in black, axon in blue), connectivity based on all axo-dendritic appositions (in red) is characterized by an extremely wide distribution of synapses per connection and almost 100% connection probability (right, pooled data from efferent connections to L23PCs of $n = 100$ L23SBCs). (Inset) A selected axon collateral with all appositions. (2) After general pruning. For the same exemplar, L23SBC, randomly removing a fraction of appositions removes the right side of the distribution of synapses per connection (right). (3) After multi-synapse pruning. Removing connections formed by too few appositions prunes the left side of the distribution of synapses (right) but leaves short inter-bouton intervals. (4) After plasticity pruning. The last step randomly removes more connections (right), leading to correct inter-bouton-intervals and connection probabilities.

(B) Examples of in silico multi-synapse connections resulting after the four-step apposition to synapse conversion algorithm. The pre- and postsynaptic m-types forming the synaptic connection are indicated. The presynaptic neuron is shown in yellow, postsynaptic neuron in black, and synaptic contacts as red circles.

mature and reach higher layers. Pooling all excitatory and inhibitory cells in each layer reveals that recurrent excitation increases with cortical depth while recurrent inhibition is weak in all layers, that descending interlaminar projections are stronger than

ascending projections, and that intralaminar inhibition is weakest in layer 4 (Figure S6C).

The seven statistical instantiations of the mean microcircuit (BioM) yield 636 ± 10 million appositions and 36.5 ± 0.5 million

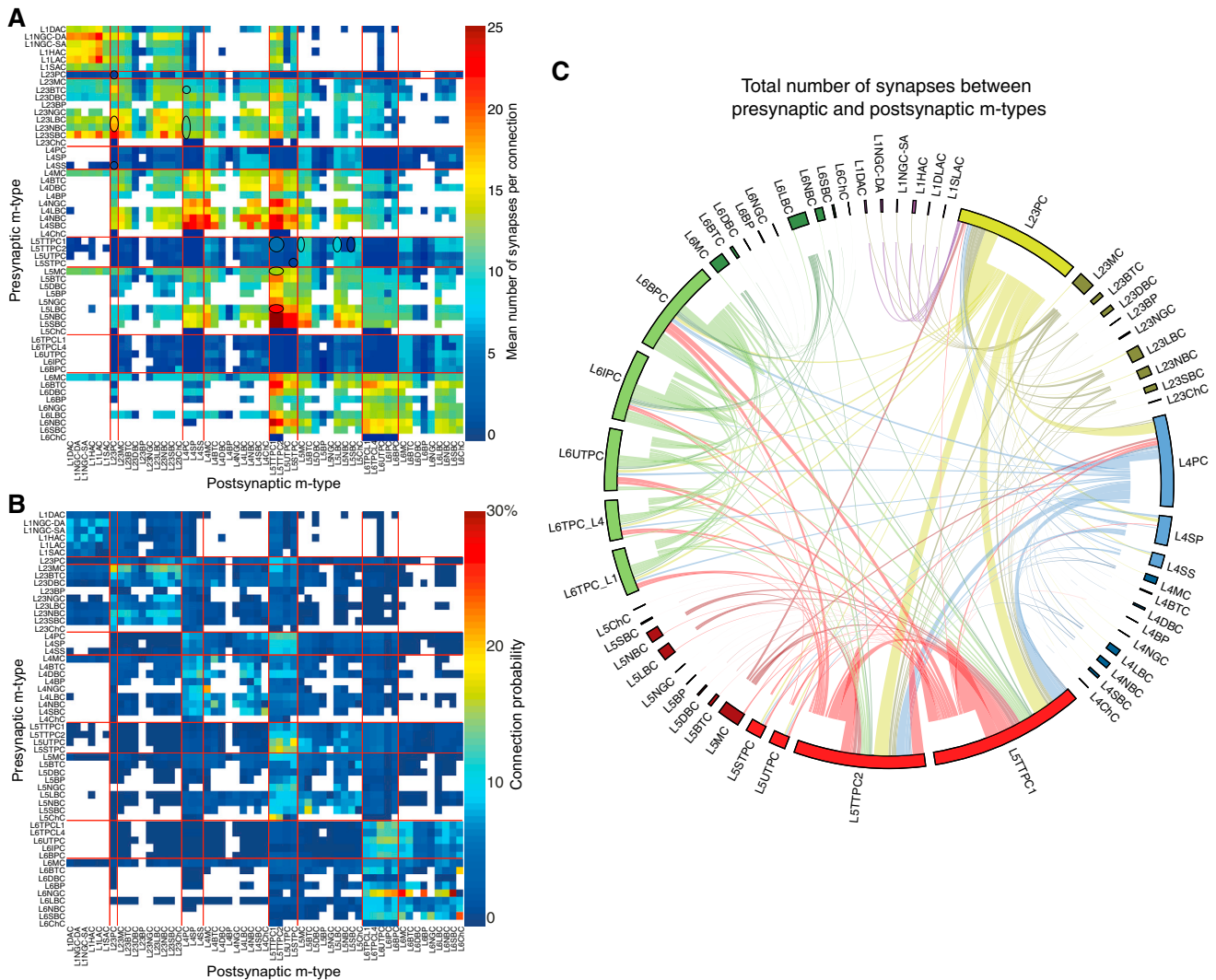


Figure 7. Predicted Synapse Numbers and Connection Probabilities

(A) Synapses per connection. A matrix of the average synapses per connection for multi-synapse connections formed between the 55 m-types (1,941 biologically viable connection types).

(B) Connection probabilities. A matrix of average connection probabilities within 100 μm .

(C) The connectome of the reconstructed microcircuit grouped by m-type (i.e., 1,941 m-type pathways). Colors group m-types by layer. Thickness of ribbon proportional to the number of synapses; inner ring segments, outputs (axons); outer ring segments, inputs (dendrites).

See also [Figure S5](#) for anatomical details of an exemplary pathway; [Figure S6](#) and [Table S1](#) for more details of synaptic innervation strength; and [Figure S20](#) for a comparison of the predicted connectome to a recent EM study. See also [Movie S1C](#).

synapses (25.8 ± 0.4 million excitatory and 10.6 ± 0.2 million inhibitory; $n = 7$; [Table S1](#) and [Movie S1C](#)). The lower variability of the statistical instantiations compared to the individual reconstructions (Bio1–Bio5; [Table S1](#)) indicates that the variation across digital reconstructions falls well within the bounds of biological variability.

From the space remaining on dendrites after accounting for predicted intrinsic connectivity (assuming 1.1 synapses/ μm ; [Datwani et al., 2002](#); [Kawaguchi et al., 2006](#); [Larkman, 1991b](#)), we predict that afferent fibers from beyond the microcircuit (extrinsic synapses) form a further 147 ± 4 million synapses (mean \pm SD; $n = 35$) ([Figures S6D](#) and [S6E](#)). The total

predicted number of synapses in the microcircuit is thus 184 ± 6 million (mean \pm SD; $n = 35$), of which only $20\% \pm 2\%$ of synapses are formed by neurons belonging to the microcircuit (i.e., intrinsic synapses), consistent with previous estimates in neocortex ([Stepanyants et al., 2009](#)). In a parallel electron microscopy study in which we determined average synapse density ($0.63 \pm 0.1/\mu\text{m}^3$; mean \pm SD; $n = 25$) and calculated the number of synapses in a comparable volume of the neocortex, we obtained 182 ± 6 million synapses. On the assumption that the average number of synapses/connection is the same for afferent fibers as for excitatory connections within the microcircuit (3.6 ± 0.04 synapses/connection; $n = 35$), we predict that

the microcircuit contains ~41 million mostly en passant afferent fibers.

The reconstructed microcircuitry reproduces numerous other experimental findings that were not used in the reconstruction process, described in a companion paper (Reimann et al., 2015). Nevertheless, it is clear that the predicted connectivity is a first draft that will be challenged and refined as experimental studies discover exceptions to the connectivity rules used here.

Reconstructing Neuronal Physiology

A series of algorithms and an automated workflow were developed to configure NEURON models to reproduce the electrophysiology of each me-type, (Druckmann et al., 2007, 2011; Hay et al., 2011) (see [Experimental Procedures](#)). In brief, we selected a morphologically reconstructed neuron and distributed Hodgkin-Huxley (HH)-type models of 13 known classes of ion channels (Figure S7) along the neuronal arbors (Figure 8A). Salient features were extracted from electrophysiological traces of e-type responses to step current pulses and data on back-propagating action potentials (Figure 8B; Larkum et al., 2001; Nevian et al., 2007). A multi-objective optimization algorithm (Druckmann et al., 2007) computed the vector of ion channel conductance densities that best reproduced features such as spike amplitudes and widths, spike frequency, and changes in frequency, and the resulting vector was transplanted into all neurons belonging to the m-type. Neurons in the resulting pool of models were challenged with a separate battery of stimuli not used to fit the vector of ion channel conductances. We then selected those that fell within observed distributions of features (~40% of models accepted; Figure 8C). This workflow provided a generic high-throughput method for modeling the electrical behavior of a potentially unlimited number of neurons of any e-type (Figure 8D). We automated the workflow to model all 207 me-types (Figure 8E), generating a pool of 121,231 unique neuron models. Exemplars can be downloaded from the NMC portal together with NEURON models of each m-type with all of their intrinsic synapses (see [Movie S2](#)). Morpho-electrical variation in the ensemble of model neurons was comparable to the biological variation observed experimentally. The quality of the final selection was quantified by comparing model and biological neurons in terms of their median z-scores for all electrical features (Figure 8E; see [Experimental Procedures](#)).

The generalization power of these models has been demonstrated previously (Druckmann et al., 2011). As a further test, we compared dendritic attenuation of synaptic potentials in the models against past experiments (Berger et al., 2001; Nevian et al., 2007). While attenuation along basal dendrites (Figure S8; space constant, $40.0 \pm 0.1 \mu\text{m}$) was consistent with these results (Nevian et al., 2007), the reconstruction displayed stronger attenuation along apical dendrites (Figure S8; $174.3 \pm 0.4 \mu\text{m}$) than previously reported ($273 \mu\text{m}$; Berger et al., 2001). However, the data in the literature were obtained from adult animals whose apical dendrites have larger diameters (Zhu, 2000) than those of the animals used in this study. In a subset of model neurons whose apical dendrites had similar diameters to those of adult animals (Zhu, 2000), attenuation was similar (Figure S8, B2, dark blue).

In most cases, transplantation of the vector of conductances to variants within the same inhibitory m-type preserved target physiology (~80% of models accepted), which was often maintained, even when conductances were transplanted to other inhibitory m-types (~60% of models accepted). This suggests that, in animals of the age used in the experiments, electrical behavior is relatively independent of the specific neuron morphologies.

Reconstructing Synaptic Physiology

To predict the physiology of the ~36 million synapses in the reconstruction, we integrated published paired-recording data and reported synaptic properties (conductances, postsynaptic potentials [EPSPs/IPSPs], latencies, rise and decay times, failures, release probabilities, etc.; see [Experimental Procedures](#) and NMC portal).

Neocortical synapses display known forms of short-term dynamics, which we used to classify synaptic connections as facilitating (E1 and I1), depressing (E2 and I2), or pseudo-linear (E3 and I3) s-types (Figures 9A and 9B) (Beierlein et al., 2003; Reyes and Sakmann, 1999; Reyes et al., 1998; Thomson and Lamy, 2007; Thomson et al., 1996; Wang et al., 2006). The s-types of specific connections were determined from the combination of their pre- and postsynaptic me-types (Ali et al., 2007; Bannister and Thomson, 2007; Beierlein and Connors, 2002; Feldmeyer et al., 2002; Frick et al., 2007; Gupta et al., 2000; Markram et al., 1998; Reyes et al., 1998; Somogyi et al., 1998; Thomson et al., 1993). Based on the available experimental data, we identified five rules to predict s-types for broad classes of connections: (1) pyramidal-to-pyramidal connections are always depressing (E2) (Feldmeyer et al., 1999; Frick et al., 2007, 2008; Gupta et al., 2000; Maffei et al., 2004; Markram et al., 1998; Mason et al., 1991; Mercer et al., 2005; Reyes et al., 1998; Thomson and Bannister, 1998; Thomson et al., 1993), (2) pyramidal-to-interneuron connections are also depressing (E2) (Angulo et al., 1999; Blatow et al., 2003; Holmgren et al., 2003; Markram et al., 1998; Reyes et al., 1998; Silberberg and Markram, 2007; Thomson and Deuchars, 1997; Wang et al., 2002), except for connections onto Martinotti, bitufted and other interneuron types displaying spike frequency accommodation, which are facilitating (E1) (Kapfer et al., 2007; Markram et al., 1998; Reyes et al., 1998; Rozov et al., 2001; Silberberg and Markram, 2007), (3) facilitation from inhibitory neurons is around two times stronger than from excitatory neurons (Gupta et al., 2000; Silberberg and Markram, 2007), (4) synaptic dynamics are preserved across layers for all me-type-specific connections, and (5) any remaining connections belong to the most common s-type (type 2; E2 or I2).

Since physiological characterization of all 27,625 unique me-type-to-me-type connections is not feasible, s-types in which experimental data were missing were specified using the rules above. Parameters for the synaptic dynamics of individual synapses were drawn from experimental distributions. In this manner, we generated a complete, albeit sparsely characterized, map of synaptic dynamics (Figure 9C). Stochasticity of synaptic transmission was modeled by extending a previously reported model (Fuhrmann et al., 2002). As an independent validation of the modeled synaptic dynamics, we compared the

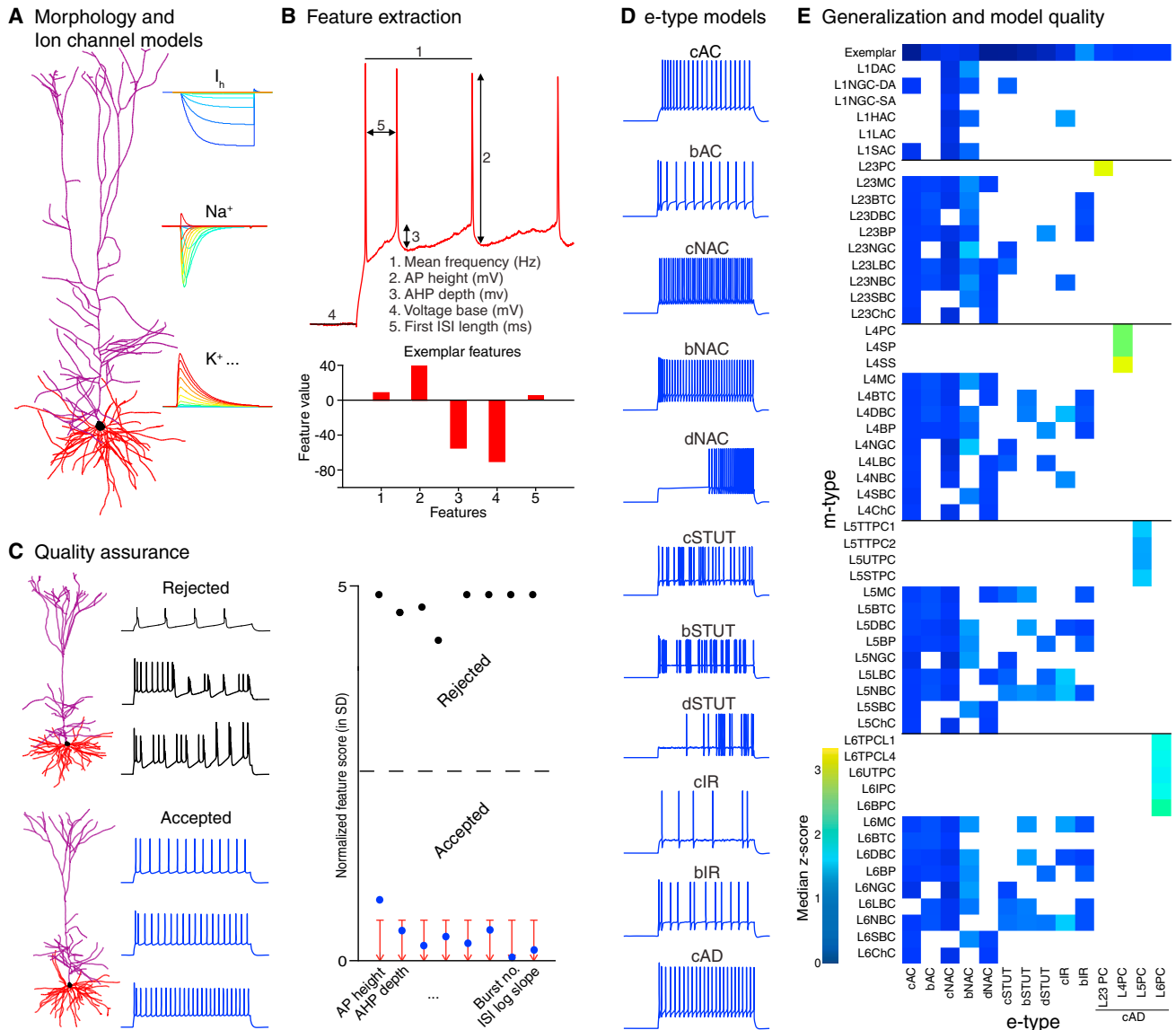


Figure 8. Workflow for High-Throughput Reconstruction of Morpho-Electrical Behaviors

(A) Morphology and ion channel models. Selection of exemplar morphology, ion channel models, and their distribution on soma, dendrites, and axon.

(B) Feature extraction. Selection of experimental traces from a population of recorded cells as targets for fitting. Extraction of voltage and spiking features from experiments.

(C) Quality assurance. Multi-objective optimization of the vector of ion channel conductance densities to match the statistics of the extracted biological features in the model. Screen out models with electrical features that do not match the statistics for equivalent features in biological recordings.

(D) Models of e-types. Shows the 11 e-types modeled.

(E) Generalization and model quality. Generalization of the vector of ion channel conductance densities to other exemplars of the same m-type; application of a standardized set of measurement protocols to each model neuron to determine generalization; quality scores for accepted models (median z-score). See also Figure S7 for properties of modeled ion channels and Figure S8 for dendritic properties.

coefficient of variation (c.v.) of first PSPs against reported experimental data ($r = 0.8$; Figure 9D; Gupta et al., 2000; Markram et al., 1998; Wang et al., 2006).

We then applied unitary synaptic conductances obtained in previous experiments that also measured somatic postsynaptic potentials (PSP) between specific pairs of m-types and compared the resulting in silico PSPs with the corresponding

in vitro PSPs (Figure 10A). The in silico PSPs were systematically lower. Since the neuron models and the numbers and locations of synapses between pairs of m-types had been validated, we hypothesized that the reported synaptic conductances had been underestimated, because of inadequate compensation for space-clamp errors (Feldmeyer et al., 2002; Gupta et al., 2000; Rinaldi et al., 2008). To quantify the underestimate,

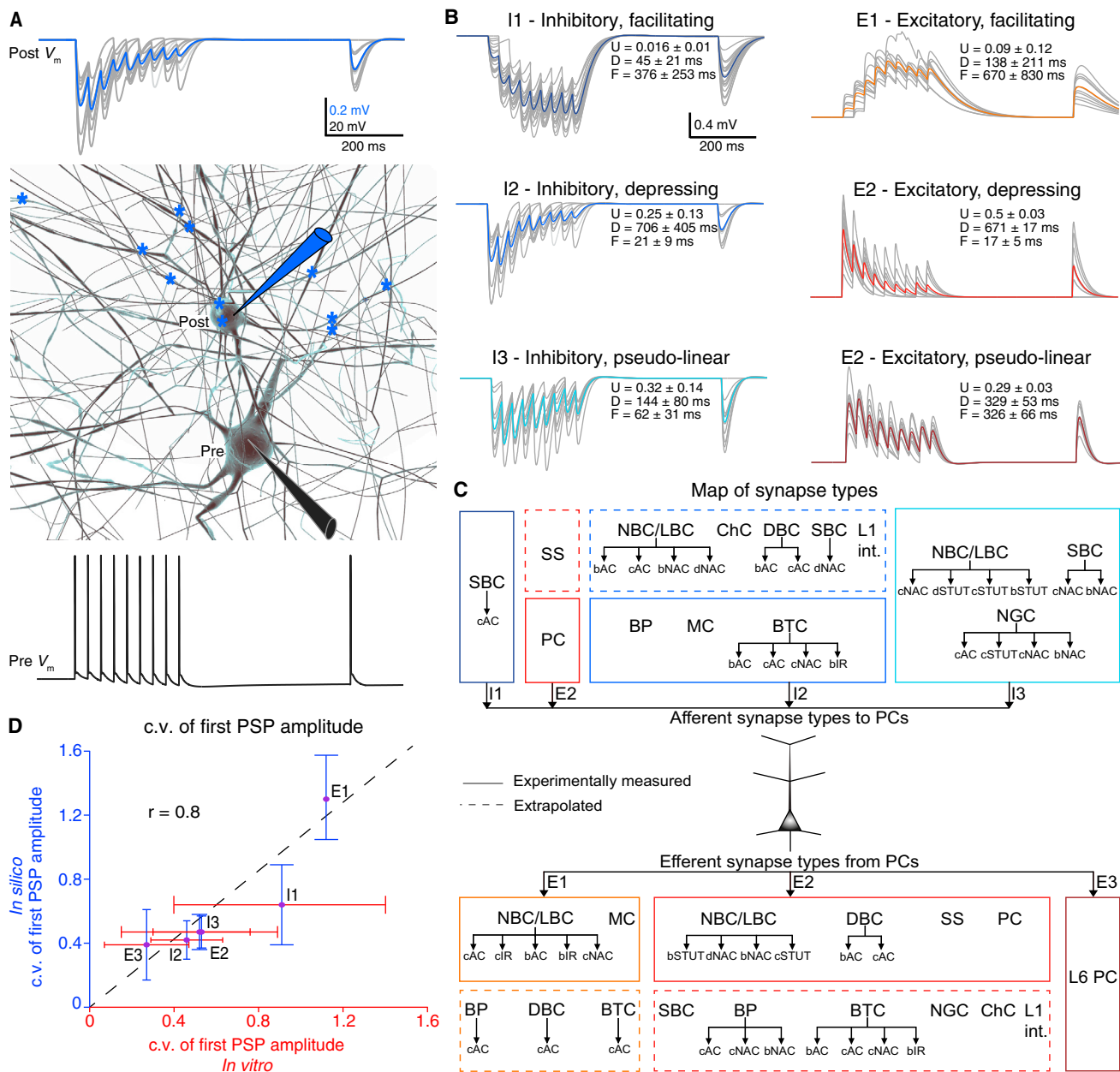


Figure 9. Reconstructing Dynamic Synaptic Transmission.

(A) In silico synaptic connection. Experimental protocol recreated in silico to obtain the frequency dependence of synaptic transmission between pairs of neurons. A presynaptic L4NBC (black pipette) was stimulated with a 30 Hz pulse train to evoke eight APs + 1 “recovery” AP (bottom trace), resulting in inhibitory depressing responses (top traces) in the postsynaptic L4SS (blue pipette); 30 individual trials in gray, average in blue. The connection was mediated by 12 synaptic contacts (blue stars).

(B) Synapse types (s-types). Parameters describing six s-types in the Tsodyks-Markram phenomenological synapse model (see [Experimental Procedures](#)).

(C) Map of predicted synaptic dynamics. Previously established mapping rules were used to constrain s-types for connections that have not yet been characterized experimentally (see [Figure S8](#)), yielding a complete map for all 1,941 m-type-to-m-type connections in the reconstructed microcircuit.

(D) Validation. Trial-to-trial variability for different s-types in silico compared to in vitro data. Dots and error bars show mean \pm SD of the data; dashed line shows regression fit.

synaptic conductances were adjusted until in silico PSPs matched experimental levels ([Figure 10B](#) and [Table S2](#); [Angulo et al., 1999](#); [Le Bé et al., 2007](#); [Feldmeyer et al., 2006](#); [Feldmeyer et al., 1999, 2002](#); [Markram et al., 1997](#); [Silberberg and Markram,](#)

[2007](#)). The results suggested that reported conductances are about 3-fold too low for excitatory connections, and 2-fold too low for inhibitory connections ([Table S2](#); [Gupta et al., 2000](#); [Rinaldi et al., 2008](#)). Other recent studies also suggest that

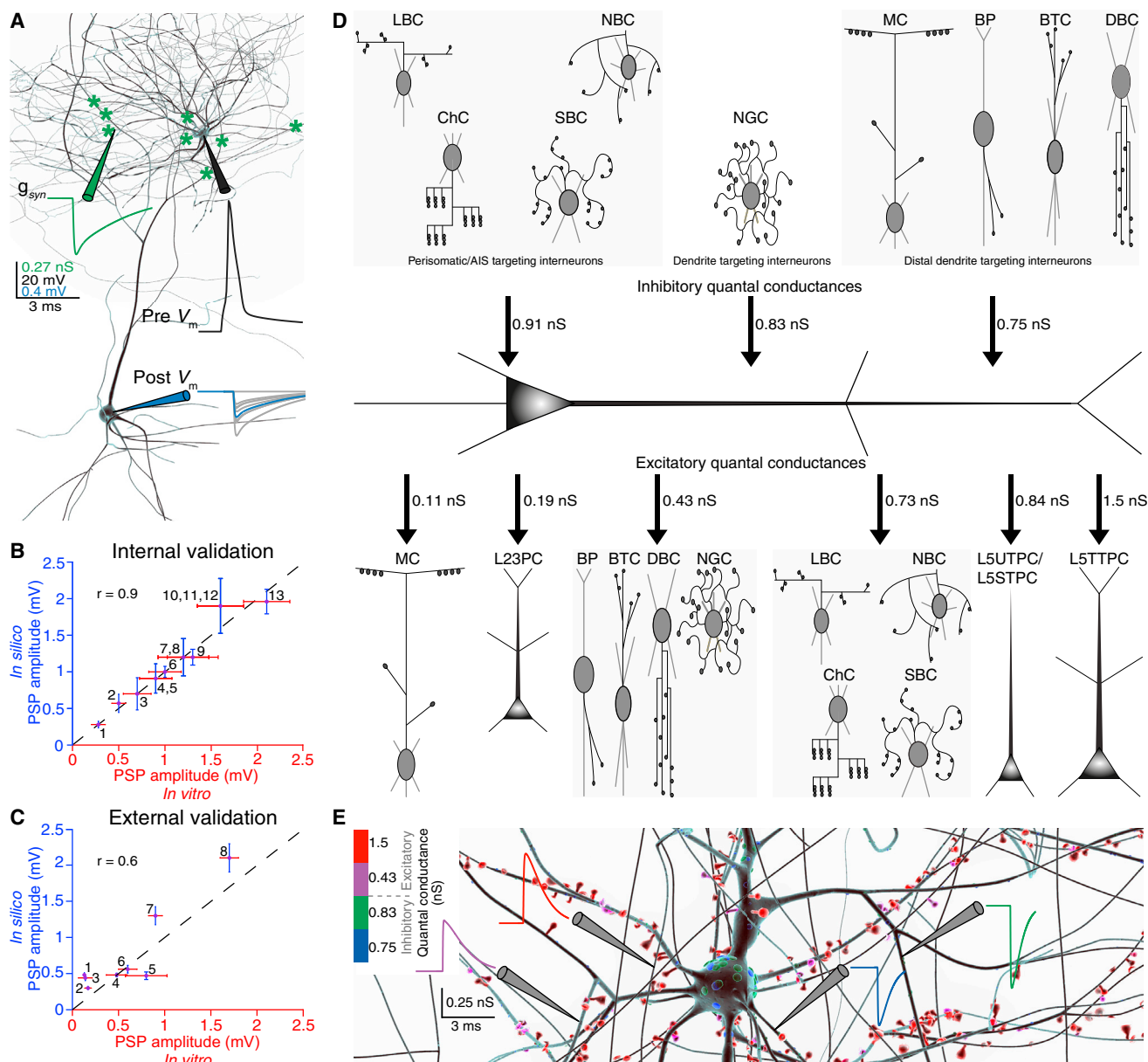


Figure 10. Reconstructing Quantal Synaptic Conductances.

(A) Unitary synaptic responses. A single AP was evoked in a presynaptic L1HAC (black pipette and trace). The postsynaptic potential was recorded at the soma of a L23PC (blue pipette; 30 individual trials in gray, average in blue); synaptic conductance was recorded simultaneously in the dendrite (green pipette and trace).

(B) Validation. Comparison of *in silico* PSP amplitudes to *in vitro* characterized connections ($n = 9$; mean \pm SD; Table S2), explicitly correcting reported conductances for space-clamp errors (see Experimental Procedures). Dots and error bars show mean \pm SD of the data; dashed line shows regression fit.

(C) Validation. As B for connections that lack conductance estimates ($n = 10$; mean \pm SD; Table S4). Conductances were generalized from B for broad classes of excitatory and inhibitory connections (see Experimental Procedures).

(D) Quantal synaptic conductances. In the absence of experimental data for postsynaptic potentials, synaptic conductances were generalized from data for similar connections, allowing the prediction of quantal synaptic conductances for all synapses on a neuron. Simultaneous recording of quantal synaptic conductances in a L5TTPC are shown in colored traces (excitatory, red to pink; inhibitory, green to blue).

(E) Predicted map of quantal conductances. Circles indicate connections used in B (black) and C (white) above. Black lines separate excitatory m-types. See also Figure S9 for examples of *in silico* synaptic patch and staining experiments and Table S2 for corrected conductances. See also Movie S2.

previously reported values are underestimated to a similar degree (Sarid et al., 2007; Williams and Mitchell, 2008).

For the vast majority of connection types, no experimental data for synaptic conductances were available. Therefore, we computed the average corrected synaptic conductances for broader classes of synaptic connections (e.g., E-E, E-I, I-I, I-E; see [Experimental Procedures](#)) and applied these conductances to all specific connections where data were missing. The resulting amplitudes of in silico PSPs were validated against experimental data for ten connection types not used in determining the conductances (regression, $r = 0.6$; [Figure 10C](#) and [Table S2](#)). The derived synaptic dynamics and quantal conductances compared well with previous reports (Feldmeyer et al., 2002; Ramaswamy et al., 2012; Silberberg and Markram, 2007; Thomson and Deuchars, 1997) (see NMC Portal). Using the same method, we generated a first prediction of mean synaptic conductances for all 1,941 m-type-to-m-type connections ([Figure 10D](#)). Unique quantal synaptic conductances for individual synapses were drawn from truncated normal distributions around these means ([Figure 10E](#); see [Experimental Procedures](#)).

We performed in silico paired recordings of all 1,941 m-type-to-m-type connections in the average microcircuit (BioM) and found results comparable to previously published paired recordings in vitro (see [Figures S9A–S9J](#) for an example; see NMC Portal). Obtaining the anatomical and physiological properties of all the intrinsic synapses formed onto and by any neuron has long been an experimental challenge (Crick, 1979). The reconstruction now allows in silico retrograde staining experiments for any neuron in the microcircuit, providing a detailed view of its presynaptic neurons and their synapses ([Figure S9K](#)). In silico anterograde staining for postsynaptic neurons is also possible. [Figure S9L](#) illustrates predicted locations of afferent synapses formed onto a L23_PC. [Figure S9M](#) shows the mean number of presynaptic (red) and postsynaptic (blue) neurons for excitatory (top) and inhibitory (bottom) m-types. Predicted input-output synapses for all 31,346 neurons in the BioM microcircuit and summary statistics for each of the 55 m-types, 11 e-types, and 207 me-types can be downloaded from the NMC Portal. The portal also provides NEURON models of each m-type, allowing simulation experiments exploring dendritic integration of m-type-specific synaptic inputs.

We found that the m-, e-, and s-types of inputs to any particular neuron were always strikingly different from those of its outputs (i.e., inputs and outputs were highly asymmetrical; see NMC Portal). The predicted average total synaptic conductance for single neurons was ~ 1000 nS (~ 750 nS excitatory and ~ 250 nS inhibitory conductance; based on all synapses in BioM). Predicted average quantal conductance was 0.85 ± 0.44 nS for excitatory synapses (corresponding to ~ 150 AMPA and ~ 20 NMDA receptors; Yoshimura et al., 1999) and 0.84 ± 0.29 nS for inhibitory synapses (corresponding to ~ 40 GABA_A receptors; Ling and Benardo, 1999). The average failure rate across all 1,941 m-type-to-m-type connections was $11.1\% \pm 14.1\%$.

Simulating Spontaneous Activity

To simulate reconstructed microcircuits at the level of detail described above, the NEURON simulator was extended to run on supercomputers ([Figure S10](#); Carnevale and Hines, 2006;

Hines and Carnevale, 1997; Hines et al., 2008a, 2011, 2011; Migliore et al., 2006), and additional functionality was developed to support in silico experimentation (see [Experimental Procedures](#)). We then used simulations to investigate the neuronal activity of the reconstructed microcircuitry ([Figure 11A](#)) under different conditions. We began by simulating spontaneous activity during tonic depolarization (see [Movie S3A](#)), attempting to mimic previous in vitro experiments (see [Experimental Procedures](#)). Under these conditions, neurons belonging to all m-types were active and the network exhibited spontaneous slow oscillatory population bursts, initiated in L5, spreading down to L6, and then up to L4 and L2/3 with secondary bursts spreading back to L6 ([Figure 11B](#)). Despite apparent global synchrony, the 55 m-types generated diverse patterns of spiking ([Figures 11C](#) and [11D](#)).

To allow comparison with the in vitro experiments, from which the physiological data were obtained, we reconstructed a virtual brain slice (a mesocircuit) that was $230 \mu\text{m}$ thick and whose width was equivalent to that of seven microcircuits (containing a total of 139,834 neurons) ([Figure 12A](#); see [Movie S3B](#) and [Experimental Procedures](#)). The virtual slice reproduced the oscillatory bursts (~ 1 Hz) found in the previous microcircuit simulations ([Figure 12B](#)), which are comparable to those found in in vitro experiments (Lőrincz et al., 2015; Sanchez-Vives and McCormick, 2000).

In vitro experiments are typically performed at $2 \text{ mM } [\text{Ca}^{2+}]_o$, while the level of $[\text{Ca}^{2+}]_o$ in vivo is reported to lie in the range $0.9\text{--}1.1 \text{ mM}$ (Amzica et al., 2002; Jones and Keep, 1988; Massimini and Amzica, 2001; also see Borst, 2010), increasing in oscillatory cycles to $1.2\text{--}1.3 \text{ mM}$ during the transition from wakefulness to sleep (Amzica et al., 2002; Heinemann et al., 1977). Although it is not possible to fully mimic in-vivo-like conditions, we nonetheless explored the behavior of the circuit at these lower Ca^{2+} levels as an approximation of the in vivo condition.

It is well known that the $[\text{Ca}^{2+}]_o$ in the extracellular space modulates the probability of neurotransmitter release (Borst, 2010; Ohana and Sakmann, 1998; Rozov et al., 2001). We therefore modified the probability of release, consistent with experimental data for the specific sensitivities of different s-types to changes in $[\text{Ca}^{2+}]_o$ (Gupta et al., 2000; Rozov et al., 2001; Silver et al., 2003; Tsodyks and Markram, 1997) ([Figure S11](#); see [Experimental Procedures](#)). We found that, in the low Ca^{2+} condition, slow oscillatory bursting disappeared and the neuronal activity became asynchronous and irregular ([Figure 12C](#)). To validate this in silico finding, we performed multi-electrode array recordings in vitro ([Figure 12D](#)) in high and low Ca^{2+} conditions (see [Experimental Procedures](#)). As predicted by the simulations, we found that the slow oscillatory bursts present in high Ca^{2+} ([Figure 12E](#)) were replaced by asynchronous and irregular activity under low Ca^{2+} conditions ([Figure 12F](#)).

We then used the virtual slice to explore the behavior of the microcircuitry for a wide range of tonic depolarization and Ca^{2+} levels. We found a spectrum of network states ranging from one extreme, where neuronal activity was largely synchronous, to another, where it was largely asynchronous (the synchronous-asynchronous [SA] spectrum; [Figure S12](#)). The spectrum was observed in virtual slices, constructed from all 35 individual instantiations of the microcircuit (seven for each of Bio1–Bio5) and all seven instantiations of the average microcircuit (BioM).

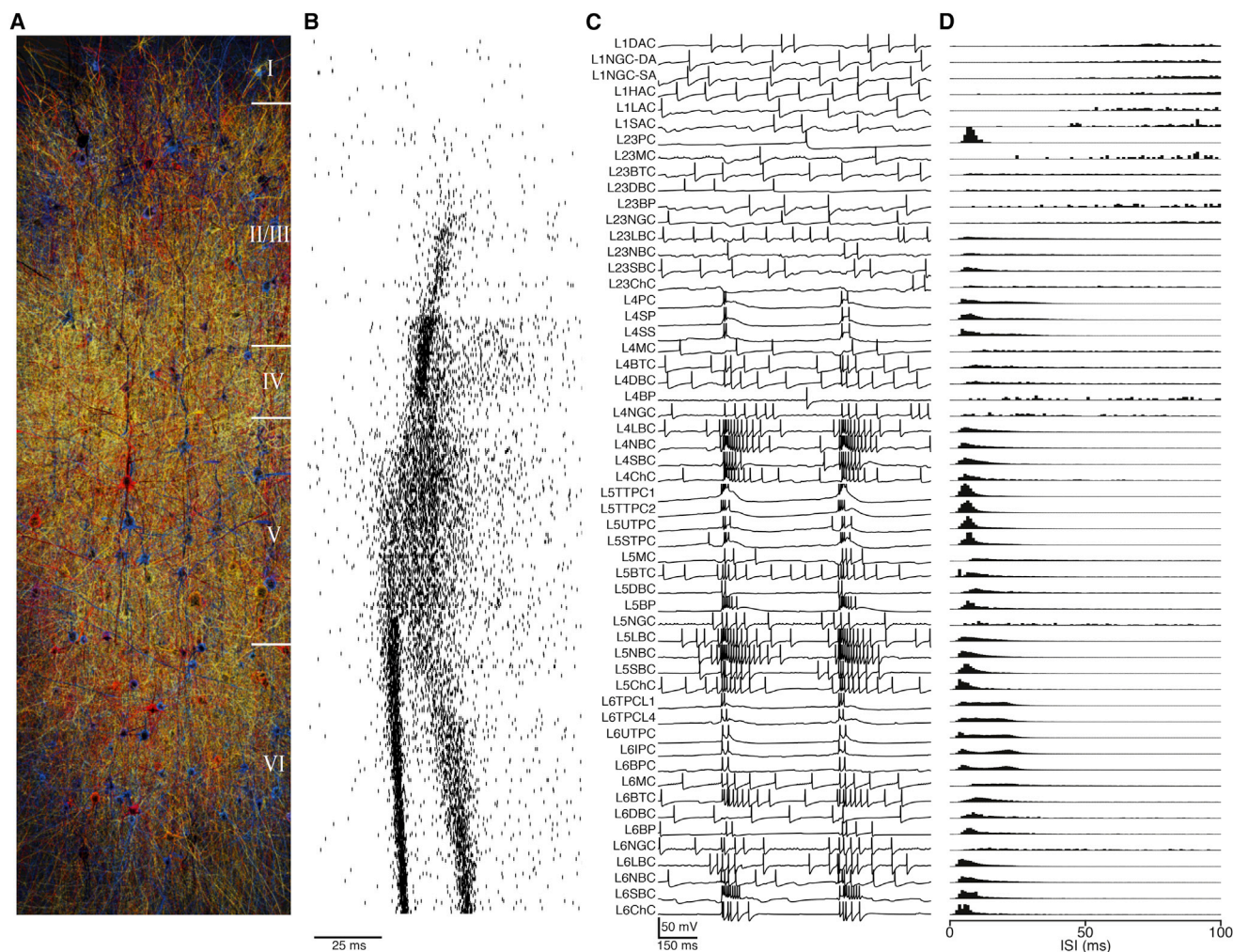


Figure 11. Simulation of the Reconstructed Microcircuit

(A) Simulation of spontaneous activity. Individual neurons at different levels of depolarization in the microcircuit are colored according to a heatmap (blue, hyperpolarized; red, depolarized; white, spike).

(B) Rastergrams of randomly selected neurons for each m-type during synchronous bursting.

(C) Exemplar voltage traces for each of the 55 m-types during spontaneous activity in the microcircuit (traces truncated at -30 mV).

(D) Inter-spike interval (ISI) distributions of each of the 55 m-types for the activity shown in C.

See also [Figure S10](#) for an overview of the software ecosystem surrounding the simulation of the microcircuit. See also [Movie S3A](#).

This implies that it is a highly reproducible phenomenon, robust to biological and statistical variations in parameters such as layer thickness, cell density, and composition; specific synaptic connectivity; and the specific dimensions of the microcircuit (see [Movie S3C](#)).

We observed that a change in $[Ca^{2+}]_o$ of < 1 mM can lead to a transition from the synchronous to the asynchronous state, revealing two distinct activity regimes ([Figure S12](#)). The level of $[Ca^{2+}]_o$ at the transition varied slightly across the different instantiations of the microcircuit (Bio1–Bio5; [Figure S13](#)).

Since the reconstructed microcircuit displays synaptically coupled assemblies comparable to those found experimentally ([Perin et al., 2011](#); [Reimann et al., 2015](#)), we also analyzed correlations in neuronal activity within these assemblies. Neuronal activity was found to be slightly more correlated within assem-

blies compared to randomly sampled neurons ([Figure S14](#)). Near the transition, a fall in $[Ca^{2+}]_o$ of just ~ 0.15 mM ([Figure S14](#)) led to a sharp decrease in correlated spiking, clearly demarcating a transition in the SA spectrum.

The mechanism underlying this sharp transition is likely to involve the differential Ca^{2+} sensitivities of inhibitory and excitatory synapse types. Indeed, we found that changing $[Ca^{2+}]_o$ from 2 mM to 1.3 mM alters the ratio between excitatory and inhibitory synaptic PSPs by a factor of ~ 3.5 , in favor of inhibition ([Figure S11](#)). This suggests the existence of a threshold level of Ca^{2+} beyond which inhibition is insufficient to prevent a supercritical state (see below).

The finding that differential sensitivity of s-types to Ca^{2+} levels determines the position of the network along the SA spectrum suggests that other mechanisms that change the

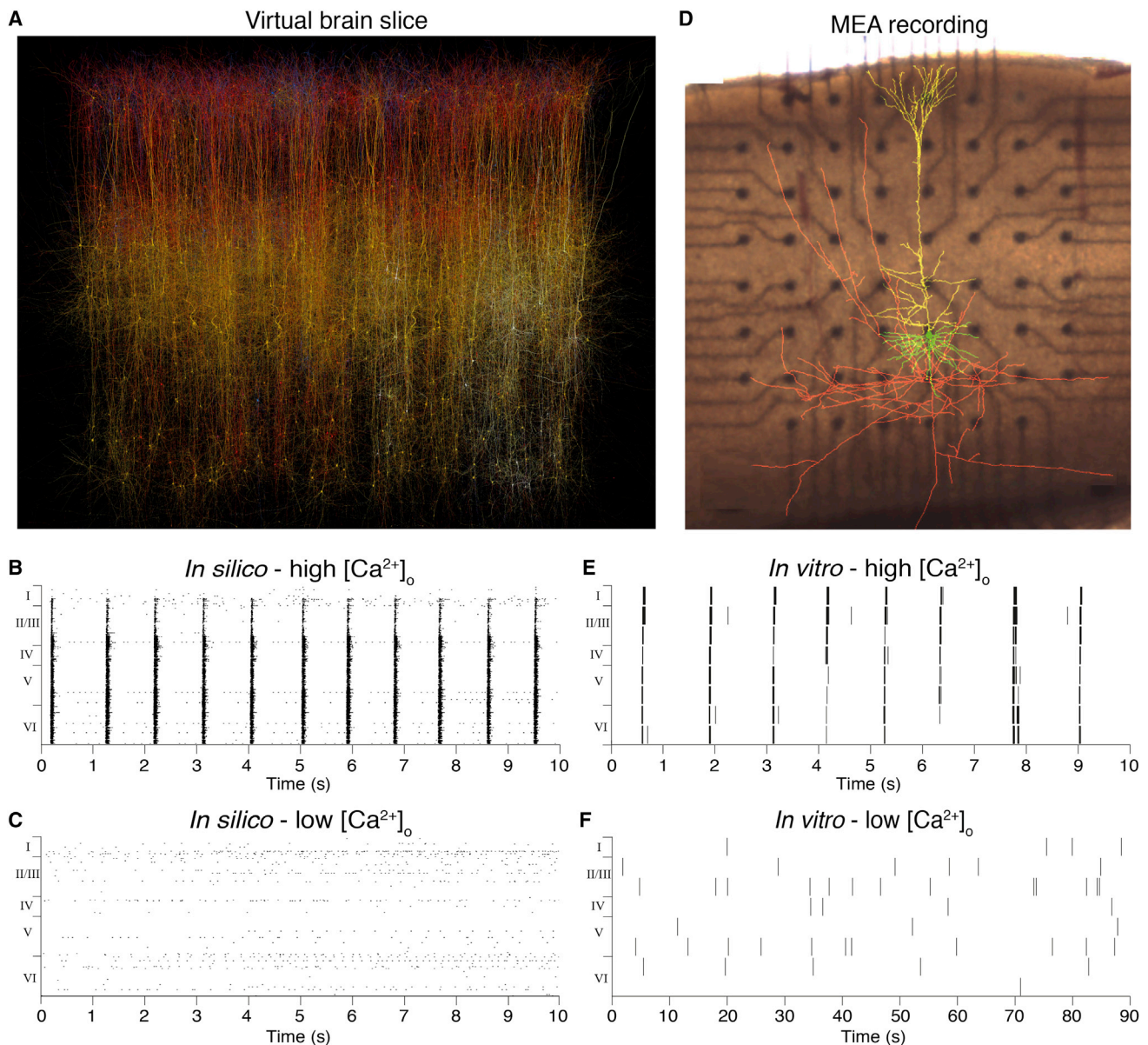


Figure 12. Predicting and Validating Synchronous and Asynchronous States in Spontaneous Activity

(A) A spontaneously active virtual slice formed from seven unitary microcircuits ($230.9 \times 2800 \times 2082 \mu\text{m}$).

(B) Rastergram of a random selection of neurons during in silico spontaneous activity under in-vitro-like conditions (somatic depolarization to $\sim 90\%$ threshold, $[\text{Ca}^{2+}]_o = 2.0 \text{ mM}$). Number of neurons displayed per layer is proportional to the total number of neurons per layer.

(C) Rastergram of a random selection of neurons during in silico spontaneous activity under in-vivo-like conditions (somatic depolarization to $\sim 90\%$ threshold, $[\text{Ca}^{2+}]_o = 1.0 \text{ mM}$).

(D) To assess network activity, 300- μm -thick cortical slices were mounted on a 3D multi-electrode array (MEA) (reconstruction of a layer 5 pyramidal cell overlaid).

(E) Experimentally observed spontaneous multi-unit activity under in vitro $[\text{Ca}^{2+}]_o$.

(F) Experimentally observed spontaneous multi-unit activity under in vivo-like $[\text{Ca}^{2+}]_o$.

See also [Figure S11](#) for $[\text{Ca}^{2+}]_o$ sensitivity of synapse types; [Figures S12](#) and [S13](#) for activity along the full spectrum of $[\text{Ca}^{2+}]_o$ concentrations and its biological variability; [Figure S14](#) for synchrony in synaptically clustered neurons along the spectrum; [Figure S15](#) for the effect of selective knockouts on microcircuit activity. See [Movie S3B](#) for a visualization of B and [Movie S3C](#) for C.

excitatory-inhibitory balance may have similar effects. We therefore performed in silico knockout experiments to understand the roles of the different layers, neurons, and connections in control-

ling the position of the microcircuit on the spectrum ([Figure S15](#)). We found that blocking activity in the upper layers tended to shift the network toward the synchronous state, while blocking the

deeper layers had the opposite effect (Figure S15A). Similarly, blocking soma-targeting basket cells produced a stronger shift toward the synchronous state than blocking other interneurons, while blocking pyramidal cells caused a shift toward the asynchronous state (Figure S15B). Corresponding differential effects were found when blocking associated inhibitory and excitatory connections (Figure S15C). These effects were observed both at high and low $[Ca^{2+}]_o$. It follows that differential regulation of layers, neurons, and connections plays an important role in controlling the position of the microcircuit along the SA spectrum, independently of $[Ca^{2+}]_o$.

Simulating Thalamic Activation of the Microcircuit

To examine spatio-temporal patterns of evoked activity, we constructed a mesocircuit consisting of a central microcircuit surrounded by six additional microcircuits. Connectivity was established for the mesocircuit as a whole, with no anatomical borders between microcircuits. An algorithm was developed to approximate input from the thalamus to the central microcircuit in such a way as to satisfy experimental constraints. We used data for the number of incoming fibers, bouton density profiles, and the numbers of synapses per connection (to layer 4) for the ventral posteromedial (VPM) thalamic input to the barrel region of somatosensory cortex (Constantinople and Bruno, 2013; Gil et al., 1999; Meyer et al., 2010b). To represent the number of fibers, we instantiated one fiber centered in each minicolumn with a horizontal spread (Meyer et al., 2010b). We then used the layer-by-layer bouton density profiles (Meyer et al., 2010b) (Figure 13A, left), experimental measurements of the mean number of synapses per thalamic connection in layer 4 (Amitai, 2001; Gil et al., 1999), and the multi-synapse principle (see above and Experimental Procedures) to predict the synapses that each thalamic fiber forms onto different m-types (Figure 13A, right). The reconstruction reproduced the number of synapses formed on L4PCs (Figure 13B) (Amitai, 2001; Gil et al., 1999) and predicted, for example, an average of ~ 12 synapses on L5 pyramidal neurons (Figure 13B), more than for L4PCs. Overall, we predicted that each thalamic fiber innervates 903 ± 66 neurons (mean \pm SD; $n = 100$ fibers; Figure 13C; 775 ± 57 excitatory and 83 ± 11 inhibitory neurons) with an average of 8.1 ± 4.2 synapses/connection. In total, we found that thalamic fibers form ~ 2 million synapses in the central microcircuit ($\sim 1\%$ of synapses across all layers; see Meyer et al., 2010b).

Thalamocortical synaptic transmission was modeled using in vitro data on synaptic dynamics (Figure 13D, left; Amitai, 2001; Gil et al., 1999) and the generalized excitatory-to-excitatory conductances derived above (i.e., similar to L4_EXC, E2 s-type; see Experimental Procedures). Ca^{2+} dependency was modeled as for other excitatory connections. The resulting synaptic transmission was validated by comparing in silico PSPs in L4 and L5 PCs in low- Ca^{2+} conditions against previous in vivo reports (Bruno and Sakmann, 2006; Constantinople and Bruno, 2013; Figure 13E). Distributions of PSPs in L4 and L5 PCs in high- Ca^{2+} conditions were also predicted (Figure 13D, right).

With $[Ca^{2+}]_o$ at 1.25 mM and moderate depolarization (in-vivo-like conditions), the main response to stimulation of thalamic fibers was in L4 to L6 (Figure S16A). Examination of the spiking activity of a random selection of neurons, covering all 55 m-types, showed that most m-types in these layers re-

sponded to the stimulus (Figure S16B). To investigate the effects of a graded stimulus, we used a single synchronous spike to activate a progressively increasing number of fibers innervating the center of the mesocircuit. With $[Ca^{2+}]_o$ at 2.0 mM and zero depolarization (in-vitro-like conditions), activating four or more fibers evoked a stereotypical high-amplitude PSTH response (>80 Hz, Figure 14A), similar to previous in vitro observations (Beierlein et al., 2002). In contrast, stimulation under in-vivo-like conditions produced graded responses, with 20–30 Hz oscillations emerging in the lower layers, particularly in L6, when higher numbers of fibers were stimulated (Figure 14B). While under in-vitro-like conditions, stimulating as few as four thalamic fibers produced all-or-none behavior, indicative of a regenerative state that spread across the whole mesocircuit (Figure 14C), under in-vivo-like conditions, the activity remained localized (Figure 14D).

With increasing Ca^{2+} levels, the stimulus response curves measured during the first 10 ms of thalamic stimulation shifted from a linear to a sharp sigmoidal shape (Figure 14E). Analysis of the velocity of spread revealed a qualitative difference between the synchronous and asynchronous regimes. In the synchronous regime, the spread of activity accelerated over time, while in the asynchronous regime, it was constant until the amplitude of the activity fell to zero (Figure 14F). This suggests that, in the synchronous regime, inhibition cannot act fast enough to curb the excitation and prevent uncontrolled spreading activity. Correlated activity is maximal in the regenerative regime and minimal in the non-regenerative regime (Figure S14), suggesting a dynamic range for correlations to emerge during information processing under in-vivo-like conditions. Taken together, the simulations predict that, at the average $[Ca^{2+}]_o$ reported in the awake state (Jones and Keep, 1988; Massimini and Amzica, 2001; Westerink et al., 1988), the neocortex will exhibit graded and spatially restricted activation, a prerequisite for the emergence of functional maps with high spatial resolution.

At a level of tonic depolarization where the network is spontaneously active in both the regenerative and non-regenerative regimes, we observed a spectrum of oscillations with lower frequencies (~ 1 Hz) in the regenerative regime and higher frequencies in the non-regenerative regime (~ 10 –20 Hz; data not shown). Maximum power was observed in layers 5–6 in the regenerative regime and in layers 2–3 in the non-regenerative regime (data not shown). This suggests that shifts along the SA spectrum contribute to the spectra of oscillatory frequencies observed in spontaneously active neocortex (see also Tan et al., 2014).

To establish a more complete demarcation between different activity regimes, we performed a series of simulations systematically exploring network state at varying levels of $[Ca^{2+}]_o$ and depolarization. An analysis of average firing rates demarcated the boundary between evoked and spontaneous activity (Figure 15A). The boundary between spontaneous regenerative and non-regenerative regimes was demarcated by the presence or absence of spontaneous bursting activity (Figure 15B). The transition between the evoked regenerative and non-regenerative regimes was determined by an analysis of the amplitude of the response to stimuli at the edge of the mesocircuit (Figure 15C). The combination of these activity maps demarcates four distinct activity regimes: evoked regenerative

A Thalamocortical innervation

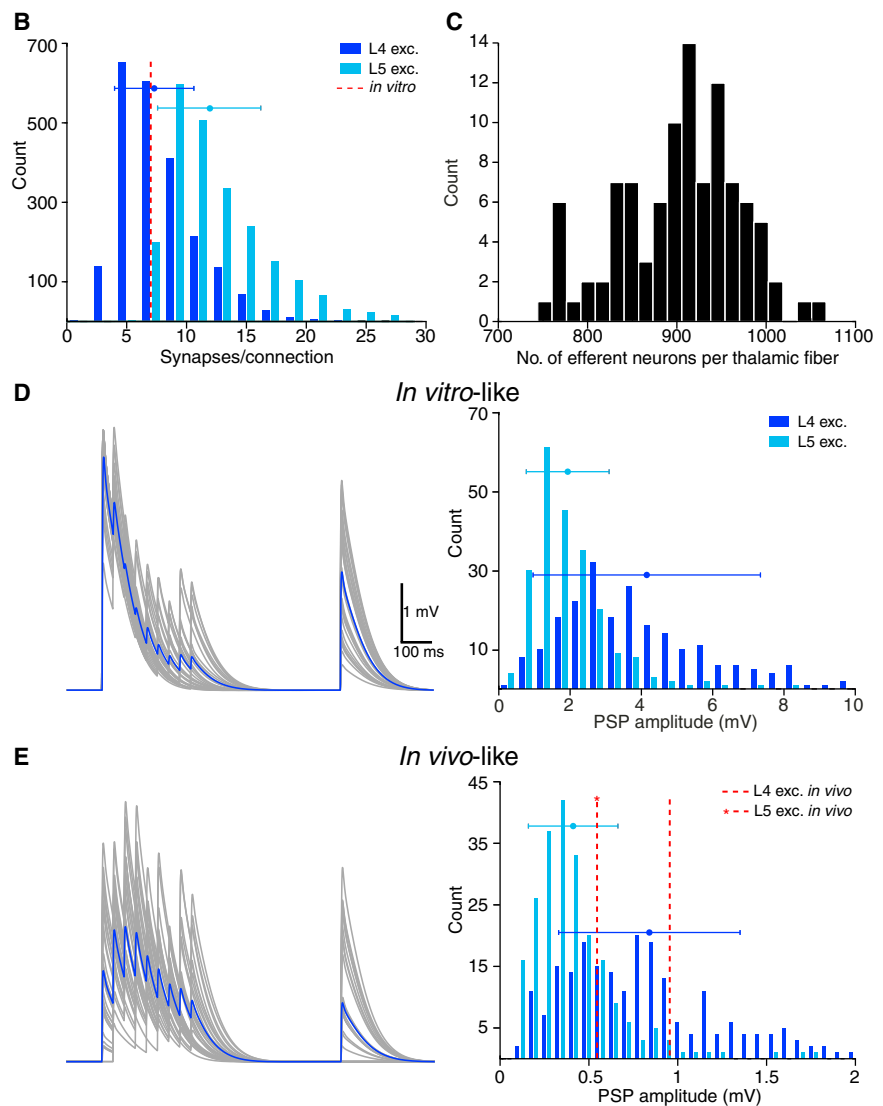
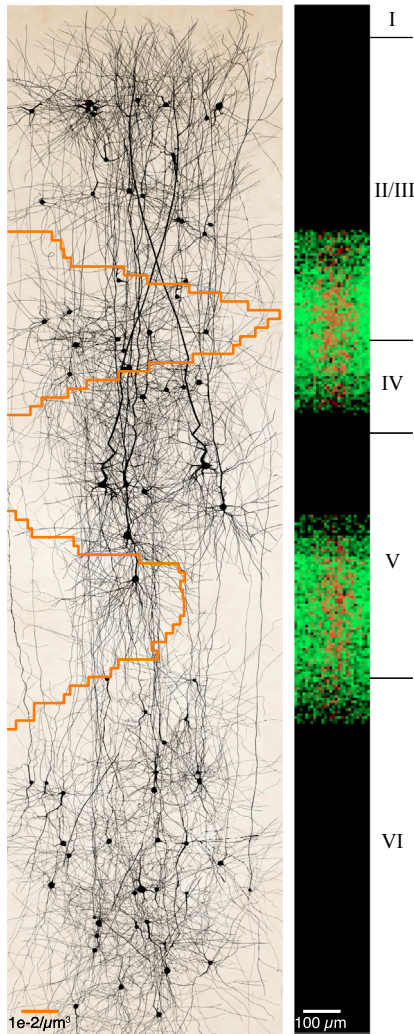


Figure 13. Reconstructing Thalamocortical Input

(A) Bouton and synapse profiles. (Left) The distribution of boutons across the depth of the microcircuit (orange line, Meyer et al., 2010b) assuming one afferent fiber from ventral posteromedial (VPM) thalamic nucleus per minicolumn, overlaid on randomly chosen neurons. (Right) The resulting synapses formed (green) with synapses formed by a single fiber (red).

(B) Distribution of the number of synapses per connection formed by the population of thalamic fibers onto L4 and L5 excitatory neurons. L4 distribution is compared against in vitro data (dashed red line; Gil et al., 1999). Horizontal bar: mean \pm SD.

(C) Postsynaptic neurons. Distribution of the number of postsynaptic neurons innervated by individual thalamic fibers.

(D) In-vitro-like conditions. (Left) Synaptic dynamics of thalamocortical connections to L4 excitatory cells (gray, 30 trials; blue, average; $[Ca^{2+}]_o = 2.0$ mM). (Right) Distribution of PSP amplitudes of thalamocortical connections to L4 and L5 excitatory cells ($[Ca^{2+}]_o = 2.0$ mM, horizontal bar, mean \pm SD).

(E) In vivo-like conditions. (Left) Synaptic dynamics of thalamocortical connections to L4 excitatory cells (gray, 30 trials; blue, average; $[Ca^{2+}]_o = 1.3$ mM). (Right) Distribution of PSP amplitudes of thalamocortical connections to L4 and L5 excitatory cells ($[Ca^{2+}]_o = 1.3$ mM, horizontal bar, mean \pm SD). L4 and L5 distributions are compared against in vivo data (L4, dashed red line, left; Bruno and Sakmann, 2006; L5, star dashed red line; Constantinople and Bruno, 2011).

See also Figure S11 for $[Ca^{2+}]_o$ sensitivity of synapse types.

(ER), spontaneous regenerative (SR), evoked non-regenerative (EN), and spontaneous non-regenerative (SN) (Figure 15D).

Reproducibility of Emergent Properties

The reconstructed microcircuitry is based on biological data from a large number of different animals and, in some cases,

from different neocortical regions that together provide statistical distributions for layer heights, neuron densities, cellular composition, and morphological and electrophysiological diversity within and across types of neuron and reflect the diversity of synaptic anatomy and physiology observed in biological experiments. The reconstruction process stochastically creates

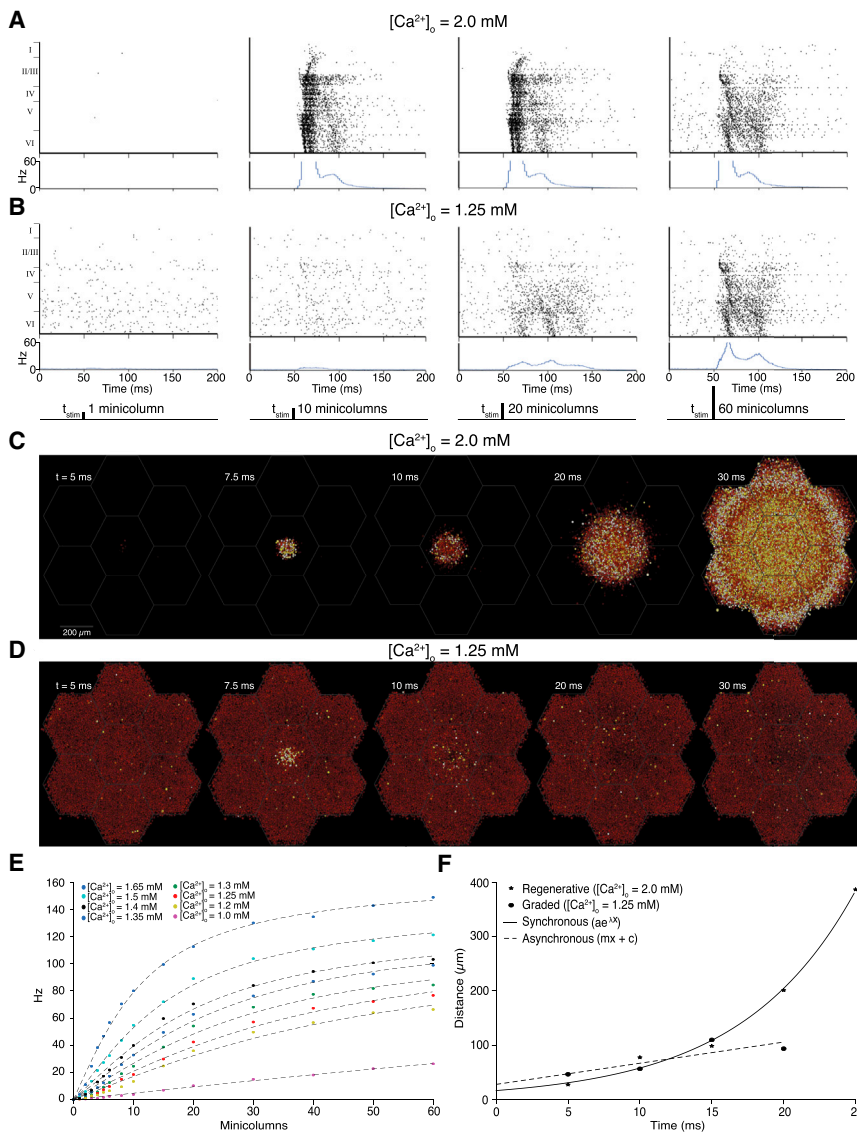


Figure 14. Activity Evoked by Thalamic Input

(A) In-vitro-like stimulus-dependent evoked activity. Raster plots (top) and PSTHs (bottom) of the response to stimulation of varying number of thalamic fibers under in vitro-like conditions (somatic depolarization to $\sim 60\%$ threshold, $[Ca^{2+}]_o = 2.0 \text{ mM}$).

(B) In-vivo-like stimulus-dependent evoked activity. Same as in A for in-vivo-like conditions (somatic depolarization to $\sim 100\%$ threshold, $[Ca^{2+}]_o = 1.25 \text{ mM}$). Stimulation times and number of fibers stimulated are shown below.

(C) In-vitro-like activity propagation. A mesocircuit under in vitro-like conditions (no somatic depolarization, $[Ca^{2+}]_o = 2.0 \text{ mM}$) stimulated with single synchronous spikes to each of 16 thalamic fibers at the center of the central microcircuit. The mesocircuit at L4 is depicted from above at different times after stimulation. Neuronal somata are rendered with a heat color map indicating level of depolarization.

(D) In-vivo-like activity propagation. Same as in D but under in vivo-like conditions (somatic depolarization to $\sim 100\%$ threshold, $[Ca^{2+}]_o = 1.25 \text{ mM}$).

(E) Stimulus response curves for various levels of Ca^{2+} and somatic depolarization to $\sim 85\%$ threshold. Response amplitude determined as the peak response for the central ten minicolumns in the first 10 ms of the response.

(F) Propagation of the wave front with time in response to thalamic stimulation (at $t = 0$), measured as the half-maximum of a one-sided Gaussian fit to the wave front. Exponential fit for the regenerative activity and linear fit for non-regenerative activity (conditions as in C and D, respectively). In the non-regenerative regime, the amplitude of the wave front was zero at 25 ms. See also Figure S16 for stimulus responses of individual m-types.

instantiations of the digital microcircuit that respect these distributions. We have previously shown that detailed synaptic physiology is largely invariant across different instantiations of the digital microcircuit (Ramaswamy et al., 2012) and that emergent parameters such as the distributions of the locations of synapses formed by different presynaptic neurons are also largely invariant (Hill et al., 2012).

To further assess the reproducibility of the reconstruction as a whole, we measured the variance of a range of its emergent anatomical and physiological properties (i.e., properties not directly specified by the data). The anatomical properties measured from seven instantiations of each microcircuit (seven instantiations of BioM and seven each for Bio1–Bio5) included total number of appositions and synapses, convergence and divergence of connectivity for each m-type, numbers of excitatory and inhibitory synapses and connections, mean numbers and types of presynaptic neurons innervating neurons belonging

to different m-types, and numbers of intra- and inter-laminar synapses and connections. In each case, we found low variance compared to the mean (see Figure S17 for a sample; see Table S1 for selected values; see also NMC Portal).

To gain a deeper understanding of the physiological variability of the digital reconstruction, we examined trial-to-trial variability in the spiking activity of individual neurons and variability across neurons of the same type, as well as variability across layers and across digital reconstructions individualized with data from five different animals (Bio1–Bio5). Cell responses to a single thalamic stimulation, roughly comparable to a single whisker deflection (stimulation of a cluster of 60 minicolumns), displayed varying degrees of trial-to-trial variability (Figure 16A). Since the digital reconstruction implements biologically grounded stochastic mechanisms for synaptic transmission, spontaneous release, and some ion channels, this was expected. However, each cell-type also displayed a characteristic delay to first spike response. In some cases, the distribution of single-neuron responses was similar to that of the population (Figure 16B, left),

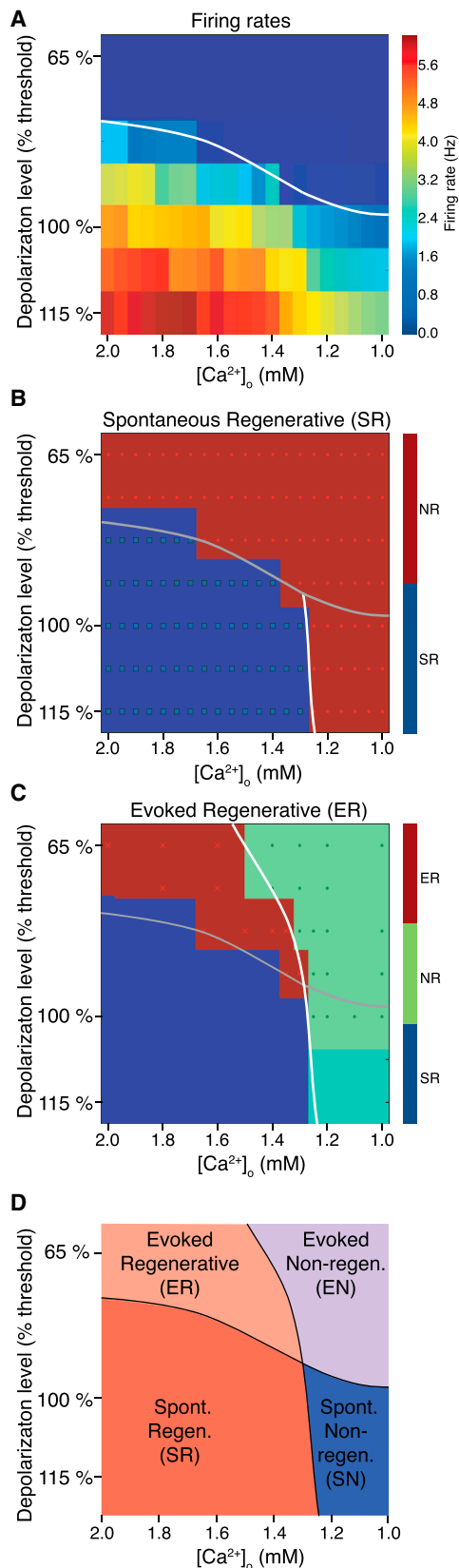


Figure 15. The Regime Map

Characterization of spontaneous and evoked activity under different levels of depolarization and $[Ca^{2+}]_o$.

(A) Average spontaneous firing rates. White line indicates interpolated transition between evoked and spontaneous regimes.

(B) Presence of spontaneous bursting activity. White line indicates transition between spontaneous regenerative (SR) and non-regenerative (NR) activity.

(C) Map of evoked regenerative (ER) and non-regenerative activity determined by the amplitude of the response to stimuli at the edge of the mesocircuit in relation to the initial response at the center. White line indicates the transition between the ER and NR regimes. Blue-green region is extrapolated to be NR. (D) Schematic map of showing the four activity regimes. Evoked regenerative, ER; spontaneous regenerative, SR; evoked non-regenerative, EN; spontaneous non-regenerative, SN.

while in others it was markedly different (Figure 16B, right). For both excitatory and inhibitory neurons, variance in response times decreased with cortical depth (Figure 16C). In all layers, trial-to-trial variability was lower than the variability between individual neurons of the same type in single trials.

Responses from neurons of the same m-type, in digital reconstructions based on data from individual animals (Bio1–Bio5), displayed higher variability across reconstructions than in different instantiations of the average microcircuit (BioM) (for an example, see Figure 16D, left). To isolate the source of this inter-individual variability, we began by re-examining the SA spectra for the reconstructions at different levels of Ca^{2+} . We found that they all displayed the spectrum but that the precise level of Ca^{2+} at the transition between the synchronous and asynchronous state was slightly different for each reconstruction, ranging from 1.23 to 1.31 mM $[Ca^{2+}]_o$ (Figure S13). We therefore repeated the simulations, setting the Ca^{2+} level such that each reconstruction was shifted to the same point along the spectrum relative to the transition. Under these normalized conditions, the variance in the responses of specific m-types across reconstructions decreased strikingly (Figure 16D, right). Figure 16E summarizes the different sources of variability for all neurons in L4 and L5.

Taken together, these results demonstrate the ability of the digital reconstruction to accommodate physiological variability while maintaining reproducibility and are evidence of its potential to generate useful biological insights. To further test this potential, we attempted to replicate results from an array of recent in vivo studies.

Reproducing In Vivo Findings

The digital reconstructions described above aimed to recreate the anatomy and physiology of an isolated slice of neocortical tissue, but not specifically to replicate any particular in vivo experiment. Nonetheless, we tested the ability of the digital reconstruction to replicate such experiments. We selected a set of recent in vivo studies in which a reasonable replication of the stimulation and analysis protocols was technically feasible. We then selected an arbitrary instantiation of BioM and used this model for all tests. In each case, we maintained the model's original parameters, without introducing modifications to fit previously reported results—a “zero tweak” strategy. All simulations were performed near the transition from the synchronous to the asynchronous state.

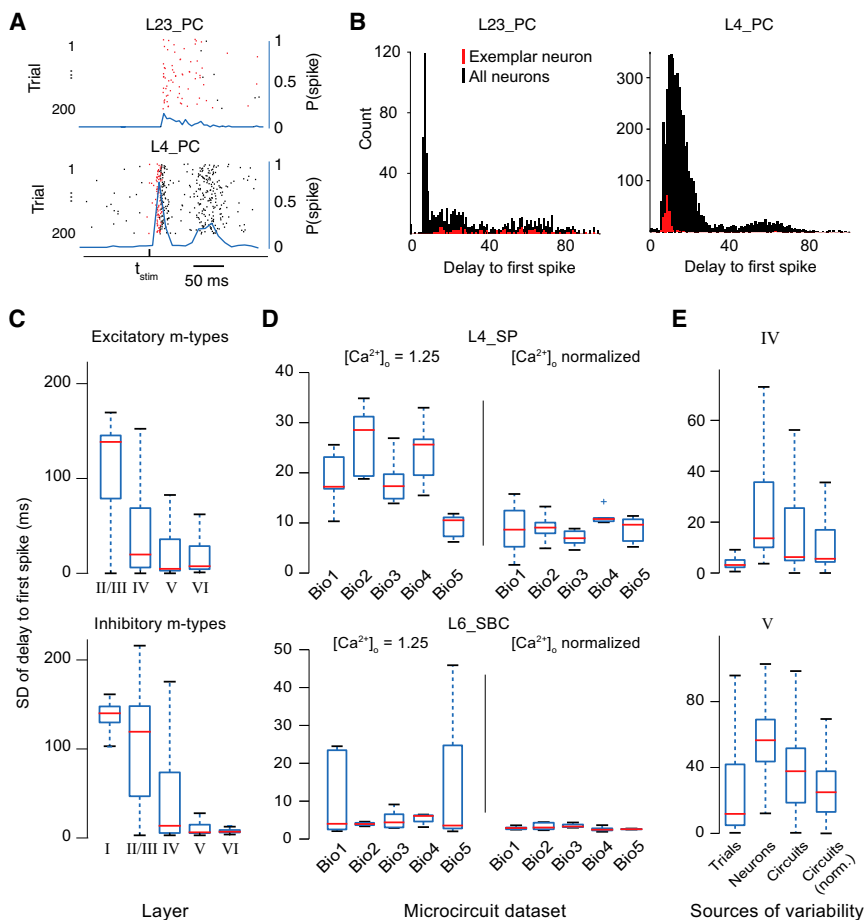


Figure 16. Reliability of Microcircuit Responses

(A) Raster plot of the spiking activity of an exemplary L23PC (top) and an exemplary L4PC (bottom) in response to simulated thalamocortical stimulation with 60 fibers. The first spike after the stimulus in each of 200 trials is indicated in red, other spikes in black. The blue line indicates the probability that the neuron fires a spike in a 5 ms bin.

(B) Histograms of the response delay (delay of the first spike after stimulus presentation) for L23PCs (left) and L4PCs (right). (Black) For 200 trials of 25 randomly chosen neurons of the indicated type. (Red) For 200 trials of the neurons indicated in A.

(C) Standard deviation of the response delay of neurons in different layers across trials. (Top) Excitatory neurons; (bottom) inhibitory neurons. (Red) For 200 trials of the neurons indicated in A. (Blue) For 200 trials of the neurons indicated in A.

(D) Standard deviation of the response delay across trials of neurons when placed in microcircuits constructed from different biological datasets (Bio1–Bio5). (Top) Of five L4SPs. (Bottom) Of five L6SBCs. (Left) Under simulated extracellular Ca^{2+} concentration of 1.25 mM. (Right) When the calcium concentration was set to a value on the border between regenerative and non-regenerative activity for that particular microcircuit. Boxes and whiskers as in C.

(E) Comparison of the different sources of variability. (Left to right) Inter-trial variability (same neuron in same microcircuit across trials); neuronal variability (same trial in same microcircuit across neurons of a given m-type); inter-circuit variability (same neuron in different microcircuit); inter-circuit variability under normalized Ca^{2+} concentrations. (Top) In layer IV; (bottom) in layer V. Boxes and whiskers as in C.

See also Figure S17 for the anatomical variability of microcircuits.

Neuronal Responses to Single-Whisker Deflection

Many *in vivo* studies of evoked neuronal activity have reported that basal activity is sparse in all cell types, that the response characteristics of individually recorded neurons display cell-type-specific diversity, and that response latencies are cell type and layer specific (Constantinople and Bruno, 2013; Reyes-Puerta et al., 2015). To test the ability of the reconstruction to reproduce these findings, we attempted to replicate some of the experiments reported in a recent study by Reyes-Puerta et al. (2015), in which the authors recorded and analyzed neuronal responses to a single-whisker deflection in the barrel cortex of anaesthetized adult rats. We approximated the stimulus as a single pulse in 60 reconstructed thalamic fibers projecting to the center of the digital microcircuit. As shown in Figure 17A1, the response to the stimulation displays cell-type-specific diversity that compares reasonably well with the results reported in Figure 3A of the Reyes-Puerta et al. study (Reyes-Puerta et al., 2015), with the exception of the OFF response, which is not as prominent. The general distribution of responses for excitatory and inhibitory cells is also comparable (Figure 17A2 *in silico* versus Figure 3B *in vivo*), though again with fewer OFF

responses. As in Reyes-Puerta et al. (Reyes-Puerta et al., 2015), most responses occurred within 10–20 ms of the stimulus and were generally led by inhibitory cells and, more specifically, by inhibitory cells in L4 and L5 (Figure 17A3 *in silico* versus Figure 4B1 *in vivo*).

Anti-correlated Inhibitory Activity Cancels Out Highly Correlated Excitatory Activity

Many previous studies have struggled to explain the uncorrelated neuronal spiking activity that is often observed *in vivo* (Celikel et al., 2004; Mazurek and Shadlen, 2002), with some suggesting that it is the result of poorly correlated excitatory activity (as expected if excitatory neurons generate a rate code), while others argue that correlations in excitatory activity are cancelled out by anti-correlated inhibition (Beierlein et al., 2000; Okun and Lampl, 2008). A model developed to address this issue by Renart et al. (2010) shows that it is indeed theoretically feasible for anti-correlated inhibitory activity to cancel out highly correlated excitatory activity (see their Figure 3). To test this hypothesis, we therefore analyzed whether this phenomenon was evident during spontaneous activity in the digital reconstruction. Although the digital reconstruction was not specifically

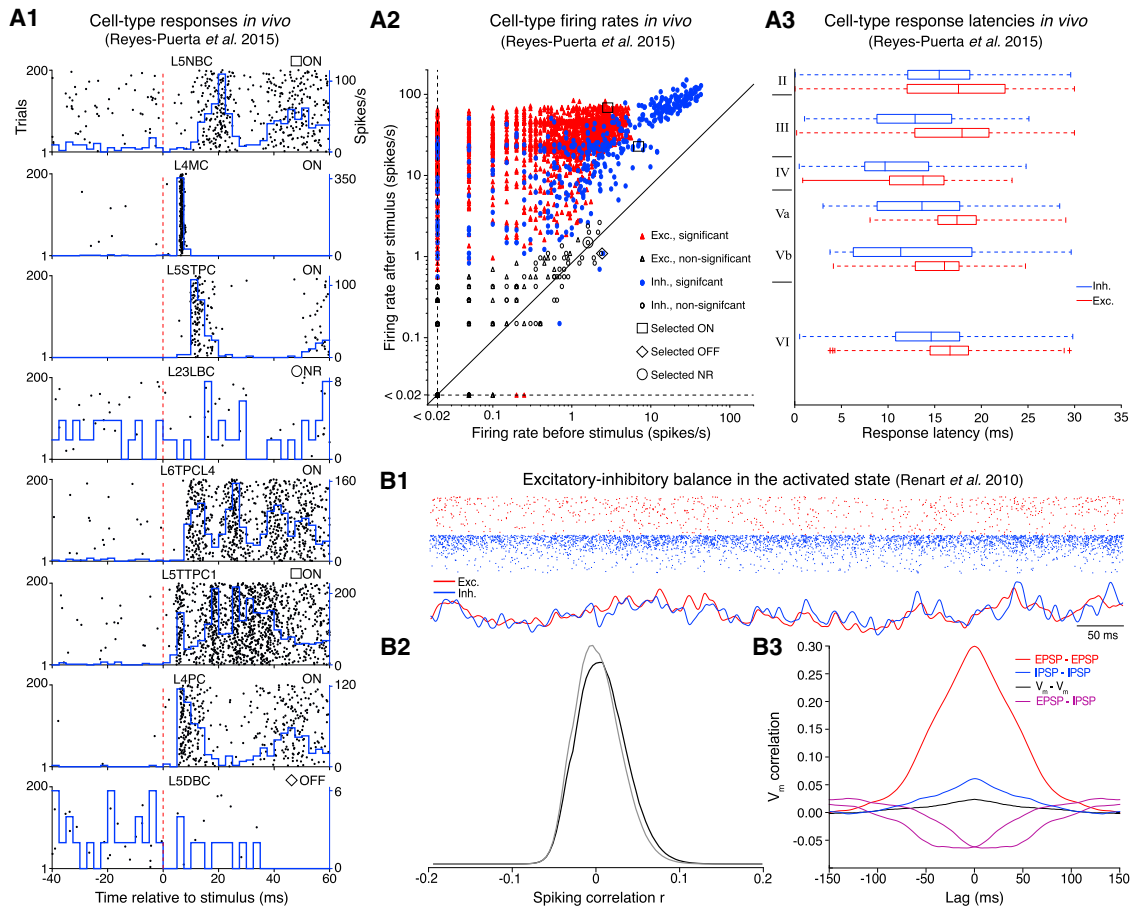


Figure 17. Cell-Type Responses In Vivo and E-I Balance

(A1) Cellular response types to simulated single-whisker deflection. Each subplot represents the activity of an individual cell, containing the raster plot aligned to simulated whisker deflection and the PSTH. Upon simulated whisker deflection, neurons increased their firing rate (ON cells), showed no change in firing rate (NR cells), or decreased their firing rate (OFF cells). (A2) Comparison of mean firing rates before and after whisker deflection plotted in logarithmic scale (2630 excitatory and 550 inhibitory neurons). Empty symbols represent neurons showing no significantly different activity in both periods (NR cells), and filled symbols represent neurons showing significantly different ($p < 0.05$) activity (ON and OFF cells). (A3) Mean first-spike latencies of inhibitory (INH) and excitatory (EXC) neurons to simulated whisker deflection, defined by first spike occurrence within 30 ms after stimulation, mean over 200 trials, for all 31,346 neurons in the stimulated column. Each box plot represents median, interquartile, and range of latencies; crosses represent outliers (2.5 times interquartile range).

(B1) Raster (top) of the spontaneous spiking activity of 500 excitatory (red) and inhibitory (blue) neurons under in-vivo-like conditions (100% depolarization and $[Ca^{2+}]_o = 1.25$ mM). Bottom curves show tracking of instantaneous population-averaged activities (transformed to z-scores, bin size 3 ms). Average firing rates of E and I cells were 1.09 ± 1.0 Hz and 6.00 ± 8.95 Hz, respectively ($n = 1,000$; mean \pm SD). (B2) Histogram of spike-spike correlations (black, count window 50 ms) and of jittered spike trains (gray, jitter \pm 500 ms). (B3) Population-averaged cross-correlograms of the somatic membrane current, when cells are held at the reversal potential of inhibition (blue) or of excitation (red) in both cells, or at one potential for one cell and at the other potential for the other cell (magenta). The black curve is for pairs at resting potential.

See also Figure S18 for details of E-I balance.

designed to produce this phenomenon, it nonetheless generated excitatory conductances in single neurons that were highly correlated but effectively cancelled out by anti-correlated inhibitory conductances (Figure 17B).

Deeper investigation revealed that spiking is correlated with momentary imbalances between excitatory and inhibitory conductances lasting <10 ms and that the timing of spikes can be predicted from the difference in the E and I conductances (Figure 17B). We also found that the precision with which these imbalances drive spiking falls dramatically as the network state shifts away from the transition in either direction. When it

shifts toward the synchronous state, the correlation is strong but broad, resulting in a temporally imprecise increase in spiking more suitable for a rate code. When it shifts toward the asynchronous regime, the correlation is sharp but too weak to effectively drive spiking, a regime more suitable for a population code based on a high degree of correlated activity (Figure S18).

Temporally Sequential Structure during Spontaneous Activity of L5 Neurons

The search for precise temporal structures in brain activity, such as synfire chains, motifs, repeated spike patterns, etc., has a

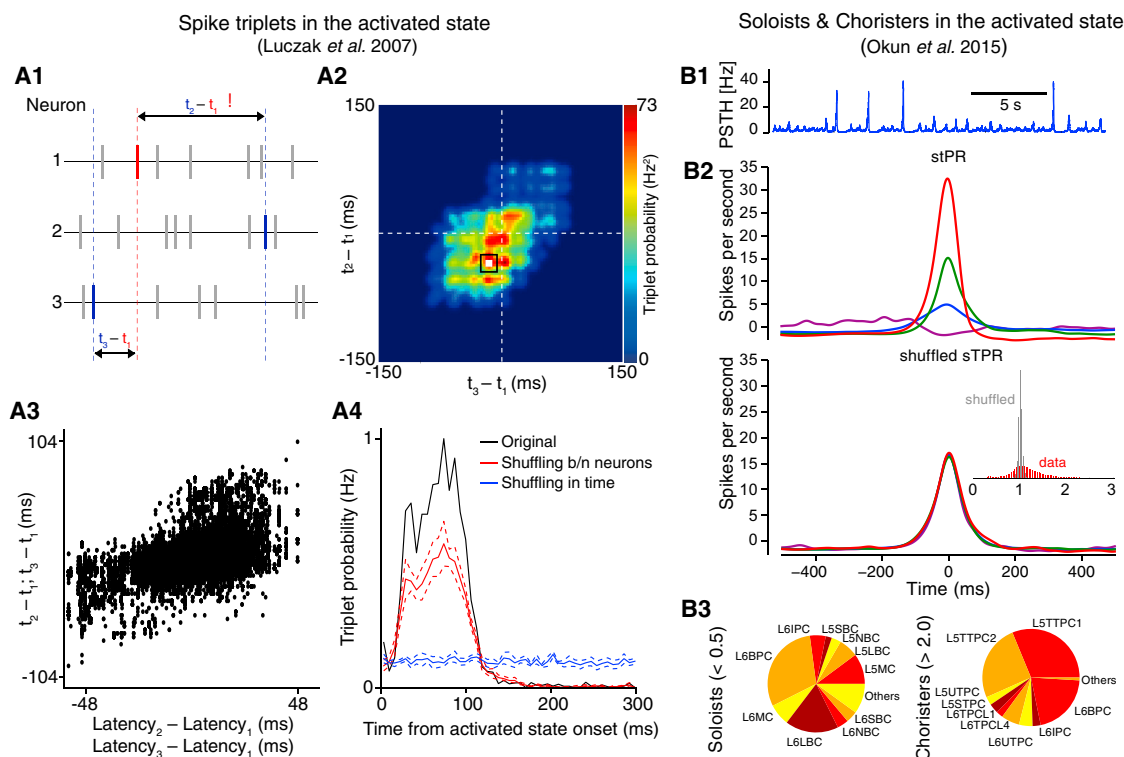


Figure 18. Triplet Structure and Diverse Population Coupling

Precisely repeating triplet structures can be predicted from individual neural latencies under synchronous stimulation (20 thalamic fibers, $[Ca^{2+}]_o = 1.25$ mM; cf. Luczak et al., 2007, Figure 5).

(A1) Schematic depicting the structure of a spike triplet for a triad of neurons. (A2) Count matrix for a representative neuron triad. Black box indicates region containing precisely repeating triplets. White square signifies mode. (A3) Correlation between neural latency differences and triplet structures. (A4) Precisely repeating triplet probability peaks shortly after onset of activated state. This peak is significant when compared with two null hypotheses (independent Poisson model, blue curve; common excitability model, red curve). Dashed lines show standard deviation.

(B1) Time course of population firing rate just below the transition to the synchronous regime (microcircuit “Bio5,” $[Ca^{2+}]_o = 1.27$ mM). (B2) (Top) Spike-triggered average of population activity (stPR) for four representative neurons in layers V and VI. (Bottom) Same as above but after shuffling (see Okun et al., 2015; cf. Figures 1E and 1G). (Inset) Distribution of the population-coupling coefficient before and after shuffling (see Okun et al., 2015). (B3) Relative fractions of m-types of soloists (population coupling < 0.5) and choristers (> 2.0).

See also Figure S19 for results under lower $[Ca^{2+}]_o$.

long history. These patterns are thought to reflect “stereotypical organized sequential spread of activation through local cortical networks,” as demonstrated recently (Luczak et al., 2007). Luczak et al. (2007) found a temporally sequential structure during spontaneous activity of L5 neurons in vivo in the somatosensory cortex (Luczak et al., 2007). In particular, they found that, after the onset of an UP state, trios of neurons generated spike motifs (triplets) with a precisely defined temporal relationship between spikes that could not be explained by random correlations during high-frequency spiking (see their Figure 5). A similar analysis of the evoked response to thalamic stimulation of L5 neurons in the digital reconstruction found the same repeating triplet structures as observed in vivo (Figure 18A). A second in silico experiment further into the asynchronous regime (i.e., at lower Ca^{2+} levels; 1.0 mM) showed no evidence of triplet structures (Figures S19A–S19C), supporting our prediction that, in the highly asynchronous regime, it is difficult for single neurons to track fine temporal structure in network activity unless the population of

presynaptic neurons becomes highly synchronized, for example, by external input.

Soloists versus Choristers

A recent study showed that some neurons in a network display spiking activity that is tightly correlated with the average activity of the population of neurons in the network (choristers), while others display a diversity of spiking patterns whose correlation with that of the population is smaller than expected by chance (soloists), suggesting that they actively avoid correlating with the rest of the population (Okun et al., 2015). We simulated the spontaneous activity of a single microcircuit in the asynchronous state but close to the transition to synchronous state for 800 s (Figure 18B1; see also Figure S19D) and analyzed the spiking activity of every individual neuron in L5 and L6 with respect to the spiking of all others. Replication of the analysis in Okun et al. (2015) yielded comparable results, although the proportion of choristers appears to be somewhat higher in the digital reconstruction (Figure 18B2 in silico versus Figures 1E and 1G in vivo).

We found that soloists are predominantly interneurons, while choristers are mainly pyramidal neurons (Figure 18B3). Pyramidal cells can be found on both extremes; they tend to be soloists when their spontaneous firing rate is high, the ratio of excitatory to inhibitory synaptic innervation is high, and most of the innervating synapses are close to their somata (data not shown).

Shifting the network further into the synchronous regime leads to an increase in the number of choristers, consistent with general recruitment of all neurons and a rate-based response. On the other hand, shifting the network further into the asynchronous regime results in a loss of both choristers and soloists during spontaneous activity (data not shown). This finding supports our prediction that, when the network is far into the asynchronous regime, single neurons cannot easily sense and respond to fine temporal structure in network activity.

Functional Implications

These replications of *in vivo* studies suggest that the digital reconstruction can yield physiologically relevant insights. We therefore went on to address two issues that it has not been possible to address experimentally, either *in vitro* or *in vivo*.

In a first experiment, we investigated the ability of single L5 pyramidal neurons to discriminate between spatially segregated inputs. As previously, we used stimuli that approximated a whisker deflection (Figure 19A). To measure how far apart the stimuli needed to be for single neurons of the microcircuit to discriminate between them, we progressively increased the spatial separation between the stimuli and measured the response of L5 pyramidal cells in terms of rate (represented by the number of spikes emitted) and timing (represented by the latency to first response). Analysis of the difference between responses yielded a measure of latency- and rate-based discrimination. Figure 19B1 shows the responses of a single, arbitrarily selected neuron to stimuli applied at locations separated by 150 μm . In this case, latency- and rate-based discrimination are both significant (Figures 19B2 and 19B3). Exploration of the discriminatory power of L5 pyramidal neurons with different separations between the stimuli and at different levels of Ca^{2+} (Figures 19C1 and 19D1) showed that many neurons discriminate between inputs separated by 150 μm or more and that a few can discriminate between stimuli with separations as small as 50 μm (i.e., approximately two minicolumns apart). In general, timing-based discrimination is much stronger than rate-based discrimination (Figures 19C2 and 19D2). Interestingly, at all separations, discrimination is strongest at Ca^{2+} levels close to the transition between the synchronous and asynchronous regimes (Figure 19D).

Unexpectedly, we noticed a spatial asymmetry in the discriminatory power of the neurons (Figures 19C1 and 19C2, shaded background). To test the reproducibility of the phenomenon, we repeated the simulation using instantiations of Bio1–Bio5 that we already knew to be highly variable (see Figure S17 and Table S1). All reconstructions showed asymmetry, but the specific degree and pattern of asymmetry was different in each case (Figure 19E1, four instantiations shown). We therefore hypothesized that the asymmetry reflects local variations in connectivity arising from the statistical instantiation of the digital microcircuit, amplified by edge effects. To test this hypothesis,

we repeated the discrimination experiment with a mesocircuit constructed as previously described ($[\text{Ca}^{2+}]_0$ 1.25 mM, separation 150 μm , see shaded background in Figures 19C1 and 19C2), taking the same microcircuit used in the previous experiment (Figures 19C1 and 19C2) as its central microcircuit (Figure 19E2, white hexagon). Under these conditions, the asymmetry was markedly reduced. We also found strong variation in overall discrimination power across the different instantiations (Figure 19E1).

In the final series of simulations, we explored the relationship between the size of the network and its emergent properties, the emergence of the transition between the synchronous and asynchronous states, and the emergence of spontaneous spatio-temporal patterns for different sized networks (10–1,000 minicolumns). In reconstructions smaller than the anatomically defined microcircuit, the transition occurred at high levels of Ca^{2+} and fell sharply with increasing size of the reconstruction, reaching a plateau in reconstructions larger than ~ 300 minicolumns (i.e., the size of the anatomically defined microcircuit; Figures 20A and 20B). Even in reconstructions as large as 1,000 minicolumns, the spectrum of states did not exhibit any further qualitative change (Figure 20B). Figures 20C and 20D show the emergence of spontaneous clustered activity as the network increases in size. We found that, in smaller reconstructions, the time course of spontaneous firing rates in different clusters of ~ 10 minicolumns was very similar and became progressively dissimilar as the reconstructions increased in size (Figure 20D). The between-cluster correlation coefficient decreased exponentially with increasing distance between clusters, also plateauing at distances comparable to the diameter of the anatomically defined microcircuit ($\sim 202 \mu\text{m}$; Figure 20E). As a related measure, we also examined the trend in correlated activity within a central set of ~ 50 minicolumns as the surrounding network increased in size. We found that, in larger networks, the correlation fell exponentially, bottoming out in microcircuits of ~ 300 minicolumns and larger ($r = \sim 0.4$; Figure 20F).

DISCUSSION

This paper presents a first-draft digital reconstruction of neocortical microcircuitry that integrates experimental measurements of neuronal morphologies, layer heights, neuronal densities, ratios of excitatory to inhibitory neurons, morphological and electro-morphological composition, and electrophysiology, as well as synaptic anatomy and physiology (see “Reconstruction Data” and Table S3). It has been validated against a spectrum of separate anatomical and physiological measurements not used in the reconstruction (see “Validation Data” and Table S3). The reconstruction provides predictions of a wide range of anatomical and physiological properties of the neocortical microcircuitry (Box 1). Simulation of the reconstruction shows a spectrum of emergent network activity states with a sharp transition from synchronous to asynchronous states. At this particular point along the spectrum, digital reconstructions reproduce a number of findings from *in vivo* studies, allowing deeper investigation of their underlying cellular and synaptic mechanisms. They also enable experiments that have not so far been possible either *in vitro* or *in vivo*. Investigation of the

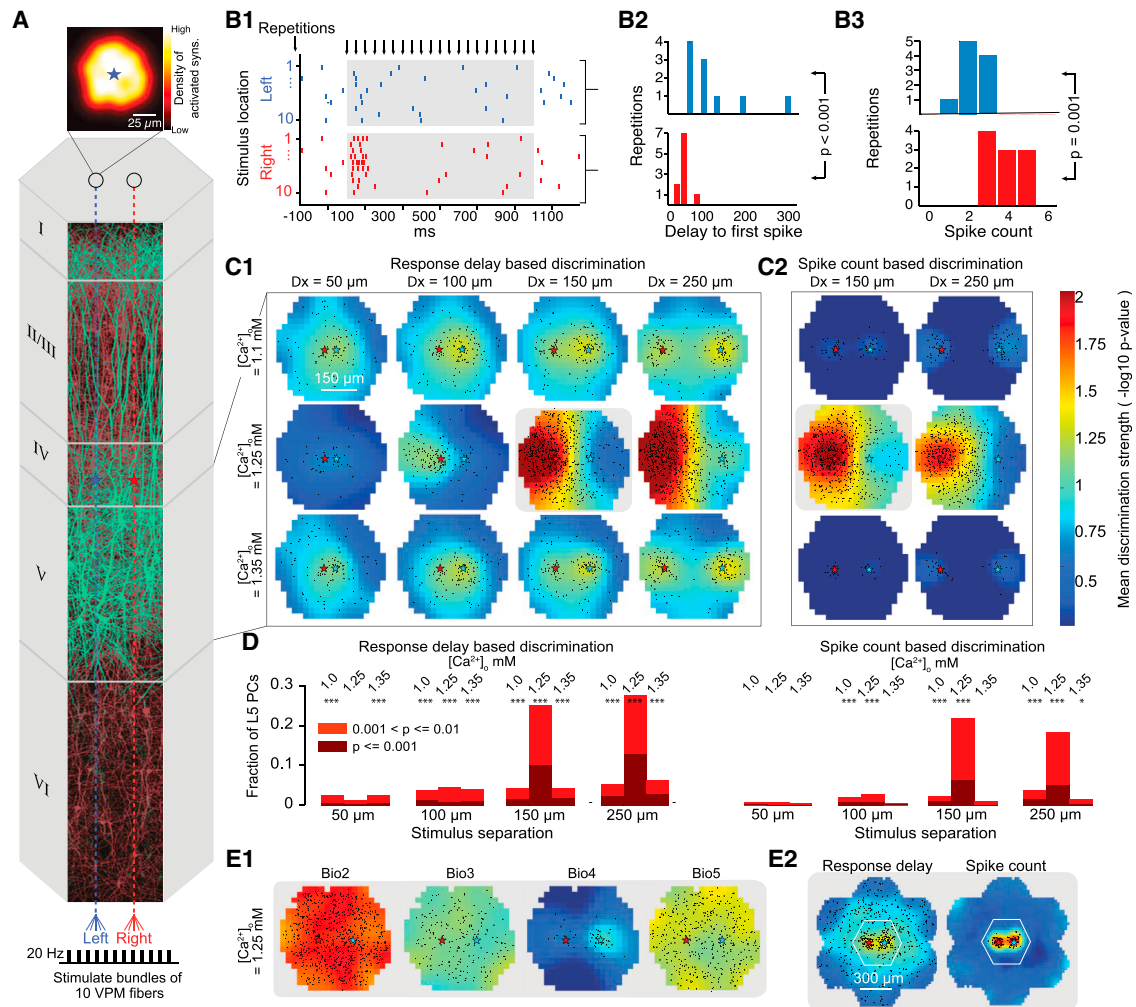


Figure 19. Spatial Resolution

(A) The ability of L5PCs in the microcircuit to discriminate between inputs given by bundles of ten thalamic fibers was examined. The stimuli were centered on locations offset from the center of the circuit to the left or right. (Top) Spatial extent of the synapses activated by ten thalamic fibers in the microcircuit.

(B1) Raster plot of spiking activity of a neuron in response to ten repetitions of spatially constrained stimuli at different locations (1 s of pulses at 20 Hz). Black arrows on top indicate individual pulses. (B2) Histogram of the delay to the first spike after the start of stimulus presentation of the neurons in B1. The difference in delay is statistically significant ($p < 0.001$, Wilcoxon rank sum test). (B3) Histogram of the number of spikes during stimulus presentation (gray window in B1). Differences in spike counts were statistically significant.

(C1) Mean discrimination strength. $-\log_{10}$ of the p value as in B2 of L5PCs at different locations is indicated as color coded. Red and blue stars indicate the centers of the two stimuli to discriminate. Black dots indicate locations of individual L5PCs with a discrimination strength >2 ($p < 0.01$). Each row indicates a different extracellular Ca^{2+} concentration. (C2) Same, for the discrimination power based on spike count as in B3.

(D) Fraction of L5PCs with a discrimination power >2 (light red) and >3 (dark red) for different conditions shown in C. Asterisks indicate instances in which the number of neurons with separation strength >2 is larger than can be explained as false positives (* $p < 0.05$; *** $p < 0.001$).

(E1) Discrimination power for a stimulus separation of $150 \mu\text{m}$ at $1.25 \text{ mM } [\text{Ca}^{2+}]_o$ for four microcircuits based on biological datasets Bio2–Bio5. (E2) Discrimination based on response delay and spike count when the same microcircuit was embedded in six surrounding microcircuits.

size of network required to reproduce key functional properties of the microcircuit shows that it is roughly equivalent to the volume of neocortical tissue used as the basis for the reconstruction. This is evidence that a network of this size is the minimum functional unit required for neocortical information processing.

Validity of the Digital Reconstruction

The reconstruction certainly includes errors due to mistakes and gaps in experimental datasets and incomplete understanding of

biological principles. For instance, additional cell type markers would improve the accuracy of the morphological composition, saturated EM reconstructions could be used to further validate the derived connectivity, more experiments reporting combined voltage and current measurements for synaptic responses will strengthen conclusions on quantal conductances and connection-specific synaptic dynamics, and further characterization of the sensitivity of different synapses to $[\text{Ca}^{2+}]_o$ may allow more accurate demarcation of the transitions between different

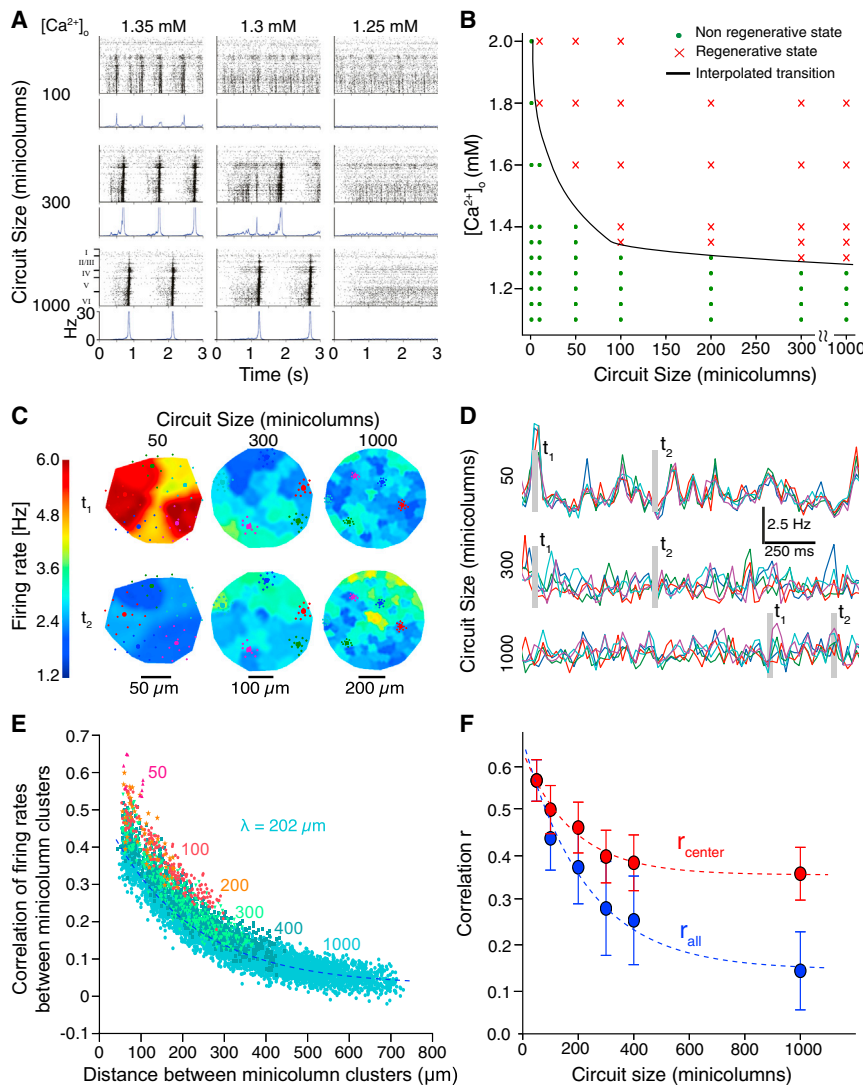


Figure 20. Emergence

(A) Transition between regenerative and non-regenerative regimes as a function of circuit size and calcium concentration. Panels of raster plots (top) and PSTHs (bottom) of spontaneous activity are shown for a selection of circuit sizes and calcium concentrations (100% depolarization).

(B) Overview of a broad range of circuit sizes and calcium concentrations as in A. Red crosses and green dots indicate regenerative and non-regenerative circuit behavior, respectively, as assessed by visual inspection. Black curve depicts interpolated transition between regenerative and non-regenerative regimes.

(C) Spatial profile of instantaneous firing rates for circuits of increasing size. Mean instantaneous firing rates were estimated for contiguous groupings (clusters) of approximately ten minicolumns using a K-means algorithm. Six spatial firing rate profiles are shown, generated by interpolating these rate estimates (see [Experimental Procedures](#)) at two selected times for three circuit sizes. Colored circles show five exemplary cluster centers.

(D) Time traces of firing rates for selected clusters. Firing rate time courses are shown for the clusters in C in corresponding color for all three circuit sizes. Dashed boxes indicate the times, t_1 and t_2 , at which spatial profiles are compared in C.

(E) Pairwise cross-correlation coefficients of cluster firing rate time courses for all cluster combinations versus inter-cluster separation for varying circuit sizes (50 to 1,000 minicolumns). Pair-wise correlation decays exponentially with distance (blue dashed line shows exponential fit to 1,000 minicolumn circuit dataset, space constant $\lambda = 202 \mu$ m).

(F) Mean pairwise cluster correlation coefficients versus circuit size for each circuit's centermost five clusters (red circles) and for all clusters (blue circles). Error bars indicate SD. Dashed curves indicate exponential fits to respective data.

activity states and may allow a more precise determination of the role played by each neuron and synapse type in maintaining and shifting regimes. For example, a recent EM study found evidence for a higher number of synapses per connection than predicted by a naive interpretation of Peters' rule ([Kasthuri et al., 2015](#)). Applying the same analysis to the digital reconstruction produced comparable findings ([Figure S20](#) in silico versus [Figures 7D, 7F, and S6B](#) in [Kasthuri et al., 2015](#)). These properties emerge in the digital reconstruction as a consequence of preferential pruning of connections with low numbers of synapses ([Reimann et al., 2015](#)).

The validation tests conducted at multiple stages of the reconstruction process reduce the risk that errors could lead to major inaccuracies in the reconstruction or in simulations of its emergent behavior. For example, validation of electrical neuron models against independent data insulates the emergent behavior of the network from the impact of our limited knowledge of ion channel kinetics and distributions. More generally, the reconstruction passed multiple tests broadly validating

its underlying anatomy and physiology. For instance, major errors in cell morphology, densities, composition, and connectivity would make it difficult to reproduce the types of neuronal assemblies discovered in 12-patch experiments, the numbers of GABAergic synapses on pyramidal somata, protein staining patterns, layer-wise synapse densities, connection probabilities, bouton densities, and distributions, etc. (see [Table S3](#)). These properties lie well within experimentally reported ranges. The reproducibility of observations and predictions in multiple reconstructions using data from different animals and incorporating statistical variations provide evidence that they are robust.

Although the reconstruction is, to our knowledge, the most detailed to date, it omits many important details of microcircuit structure and function, such as gap junctions, receptors, glia, vasculature, neuromodulation, plasticity, and homeostasis. Furthermore, it represents a snapshot of just one brain region, in one strain of male rat, at a young age. This limits the generality of the conclusions that can be drawn. For instance, in

Box 1. Microcircuit Predictions

1. The cellular composition of the microcircuit
2. Total lengths of dendrites and local axons
3. Increase in neuronal diversity with cortical depth
4. Total number of appositions and synapses
5. Total number of connections and connection types
6. Number of connections and synapses per connection between different neuron types
7. Number of connections and synapses formed by incoming fibers
8. Increase in the E-I neuronal fraction with cortical depth
9. All input and output synapses for all neuron types
10. Quantal synaptic conductances for all intrinsic synapses
11. Total excitatory and inhibitory conductances for all neuron types
12. Number and combination of pre- and postsynaptic neurons for all neuron types
13. Detailed synaptic physiology for connections between all neuron pairs
14. E-I ratios within and across layers
15. A spectrum of network states ranging from synchronous to asynchronous activity
16. Extracellular calcium regulates the network state through differential effects on synaptic dynamics
17. Role of layers, neuron, and connection types in modulating network states
18. The in vivo phenomena examined only emerge near the transition between synchronous and asynchronous states

animals of the age used for the study, dendritic morphologies have already matured to adult levels (Larkman, 1991a; Romand et al., 2011), but the ascending axons may not be fully represented and are certainly not completely mature (Romand et al., 2011). However, studies at a greater level of biological detail (e.g., including glia, receptors, and signaling pathways) and investigations of different brain regions in animals of different ages, gender, and species, as well as in disease models can use the reconstruction as a reference point. Findings consistent with the reconstruction would indicate the sufficiency of the principles of organization used in the reconstruction process; discrepancies may point to new principles. For example, if application of the connectivity algorithm to another brain region or to animals at a different age or belonging to a different species failed to yield results consistent with experimental findings, this would point to specific variations in the connectivity rules.

Failure in validation could also indicate errors in experimental data. For instance, the reconstruction indicated that cell densities from a dozen previous studies were all too low to account for spine and synapse densities, suggesting new experiments, which verified this prediction. The reconstruction also revealed that many experiments underestimate synaptic conductances and suggests that in vitro experiments that do not account for calcium level in the bath may misinterpret the relevance of their findings for in vivo conditions. These examples illustrate how the reconstruction process does not take experimental data at face value but uses complementary, related datasets to constrain the use as parameters, wherever possible.

Functional Implications

Simulations of the spontaneous and evoked activity that accounted for the differential sensitivity to Ca^{2+} of different types of synapses and that explored changes in Ca^{2+} levels revealed

a spectrum of activity states ranging from synchronous to asynchronous behavior. Varying the Ca^{2+} level profoundly changes the overall E-I balance and hence the position of the network along the spectrum, leading to a sharp transition between activity regimes. These in silico predictions were verified by new in vitro experiments.

Further simulations showed that the level of Ca^{2+} , where the transition occurred, varies across digital reconstructions that use data from different animals and that this accounts for a significant proportion of the variance in neuronal spiking and the spatial resolution of the network. We also found that a small adjustment in Ca^{2+} levels (~ 0.05 mM) in individual reconstructions significantly reduces their physiological variability. These simulations provide an example of how variations in individual neuroanatomy may lead to functional differences.

Inspired by this finding, we performed further simulations, which demonstrated that activating or inhibiting specific layers, neurons, and synaptic connections also shifts the network along the spectrum. While it is well known from previous theoretical findings that changing E-I balance changes the state of the network (Brunel, 2000; van Vreeswijk and Sompolinsky, 1996), the simulations further suggest that any mechanism that differentially changes the synaptic dynamics of different types of synapses (e.g., through neuromodulation; for reviews, see Lee and Dan, 2012; Zagha and McCormick, 2014) could alter the boundaries between activity regimes in complex ways. We speculate that other emergent properties, such as UP and DOWN states with two meta-stable fixed points, as observed in vivo (Steriade et al., 1993), which are not reproduced by the digital reconstruction, may require thalamo-cortical interactions (Hughes et al., 2002), cortico-cortical interactions (Timofeev et al., 2000), intrinsic oscillators (Lőrincz et al., 2015; Sanchez-Vives and McCormick, 2000), or neuromodulation (Constantinople and Bruno, 2011; Lőrincz et al., 2015; Sigalas et al., 2015). Modulation of cellular or synaptic

physiology may therefore serve as mechanisms to dynamically reconfigure the network to satisfy different computational requirements.

Reproducing In Vivo Findings

Although the digital reconstruction was largely based on in vitro data and was not designed to reproduce any particular experiment or to capture complex in vivo conditions, it yielded results that were qualitatively comparable to a number of major in vivo findings and made predictions beyond what was possible in these experiments, without tweaking any of the model parameters.

For example, the digital reconstruction made it possible to address a long-standing question concerning the mechanisms underlying the uncorrelated activity frequently observed in in vivo experiments (Haider et al., 2006). Previous theoretical work has shown that uncorrelated activity could be the result of tightly correlated excitatory conductances that are effectively cancelled out by anti-correlated inhibitory conductances (Renart et al., 2010; van Vreeswijk and Sompolinsky, 1996). Our simulations, using a model not specifically designed to address this question, confirm this effect as an emergent property of the network. The simulations further suggest that cortical activity in vivo approaches a critical transition along the synchronous asynchronous spectrum, beyond which regenerative activity leads to neuronal avalanches (see also Beggs and Plenz, 2003). Around this transition, spiking activity is highly correlated with fine temporal structure in synaptic input, reflected in brief moments of imbalance between excitatory and inhibitory conductances. Maximal discrimination between spatially segregated inputs, the generation of fine temporal structures such as triplets, and soloist-like and chorister-like behavior all emerge close to the transition. A recent study has experimentally characterized the plasticity mechanisms for maintaining the network close to this transition (Delattre et al., 2015).

Reproducing these in vivo findings was surprising because the digital reconstruction was based on data and architectural principles obtained from the immature rat somatosensory cortex, while many of the in vivo findings came from different neocortical regions in adult animals, sometimes belonging to other species. The fact that the reconstruction reproduces these phenomena suggests that they arise from fundamental properties of the neocortical microcircuit.

Concluding Remarks

This study demonstrates that it is possible, in principle, to reconstruct an integrated view of the structure and function of neocortical microcircuitry, using sparse, complementary datasets to predict biological parameters that have not been measured experimentally. Although the current digital reconstruction can already be used to gain insights into the way the microcircuitry operates, it is only a first step. To facilitate integration of new experimental data and challenges to the principles on which it is based, we have created a public web resource, which provides access to experimental data, models, and tools used in the reconstruction (The Neocortical Microcircuit Collaboration [NMC] Portal, <https://bbp.epfl.ch/nmc-portal>; Ramaswamy et al., 2015). This will allow the community to integrate their

own data, perform their own analyses, and test their own hypotheses.

EXPERIMENTAL PROCEDURES

A detailed description is available in the [Supplemental Experimental Procedures](#).

Data Acquisition

Neuron Morphology

Neuron morphologies were obtained from digital 3D reconstructions of biocytin-stained neurons from juvenile rat hind-limb somatosensory cortex, following whole-cell patch-clamp recordings in 300- μ m-thick brain slices (Markram et al., 1997). In some of the reconstructed neurons, bouton locations were annotated on the axon (Wang et al., 2002). Reconstruction used the Neurolucida system (MicroBrightField).

Neuron Electrophysiology

Neurons were stimulated with a set of previously described protocols (Le Bé et al., 2007; Wang et al., 2002, 2004). A subset of these stimuli was used to generate neuron models; a different subset was used to validate the models.

Synaptic Anatomy

Data on the anatomy of synaptic connections were collected from previous studies in which synaptically coupled neurons were digitally reconstructed, and putative synapses were identified using criteria identifiable in light microscopy and validated using EM. In brief, putative synapses were identified at appositions between arbors, where a bouton was also present on the axon of the presynaptic neuron (Markram et al., 1997).

Synaptic Physiology

Presynaptic neurons were stimulated with a set of previously described protocols (Gupta et al., 2000; Markram et al., 1998; Tsodyks and Markram, 1997; Wang et al., 2002, 2006). The synaptic parameters required to model the synapses were obtained by fitting the responses against the Tsodyks-Markram model for dynamic synaptic transmission (Fuhrmann et al., 2002; Tsodyks and Markram, 1997).

Tissue Immunohistochemistry

Standard immunohistochemical methods were used to label markers of cell types (Lefort et al., 2009). Stained cells were counted under light microscopy. Layer boundaries and densities per layer were computed on slices using optical dissectors on NeuN-stained tissue (West and Gundersen, 1990; Williams and Rakic, 1988) and Stereo Investigator software (StereoInvestigator 7.0, MicroBright Field). Data for each cortical layer (I, II, III, IV, Va, Vb, VI) were collected from different animals ($n = 5$). Final values for neuronal densities and layer thicknesses were corrected for shrinkage. E/I ratios were determined by soma counting in confocal microscopy imaging of dual NeuN- and GABA-stained tissue.

Electron Microscopy

Serial EM stacks were obtained for blocks of neocortical tissue, as previously described (Denk and Horstmann, 2004).

Multi-electrode Array Experiments

A 3D multi-electrode array with 60 pyramidal platinum electrodes (Qwane Bioscience SA) was used to obtain extracellular recordings from neurons in slices, as previously described (Delattre et al., 2015; Rinaldi et al., 2008). Experimental data analysis was performed in Matlab (The MathWorks) with custom scripts. Extra-cellular spikes were detected when the recorded signal crossed a dynamic threshold.

Manipulating $[Ca^{2+}]_o$

Extracellular Ca^{2+} concentration ($[Ca^{2+}]_o$) was changed by bath perfusion with artificial extracellular fluid containing a modified $[Ca^{2+}]_o$. Bath changing times were minimized by employing a pipette to remove the recording chamber solution prior to changing the subsequent solution.

Reconstruction Process

Digital Neuron Morphologies

Following 3D reconstruction, the cut ends of neuronal morphologies were restored using a repair algorithm (Anwar et al., 2009). Neuronal arbors

were digitally unraveled to compensate for tortuosity caused by shrinkage, and neuron morphologies were cloned (see [Supplemental Experimental Procedures](#)).

Electrical Neuron Models

Multicompartmental conductance-based models of neurons were generated using up to 13 active ion channel types and a model of intracellular Ca^{2+} dynamics. Axon initial segments (AIS), somata, basal dendrites, and apical dendrites were separated. Interneurons contained only one dendritic region. Each region received a separate set of channels (see NMC portal, <https://bbp.epfl.ch/nmc-portal>; Ramaswamy et al., 2015). Of the axon, only the AIS was simulated. Each AIS was represented by two fixed-length sections, each with a length of 30 μm . AIS diameters were obtained from the reconstructed morphology used for model fitting. Action potentials detected in the AIS were sent to the postsynaptic synapses with a delay corresponding to the axonal length, assuming an axonal velocity of 0.3 m/s. Neuron models were fitted using a feature-based multi-objective optimization method, as previously described (Druckmann et al., 2007).

The Microcircuit Volume

Layer thicknesses and the diameter of the microcircuit were used to construct a virtual hexagonal prism (see main text). A virtual slice was generated from a 1×7 mosaic of microcircuits as a sheet ($230.9 \times 2800 \mu\text{m}$). A meso-circuit was also generated. The meso-circuit consisted of a single microcircuit surrounded by additional microcircuits on all faces.

Cellular Composition

Cell density measurements and experimentally determined fractions of m- and me-types were used to generate the position of each cell in the volume of tissue, using E:I ratios to correct for sampling bias. Each cell was assigned the optimal morphology for its location in the volume (see [Supplemental Experimental Procedures](#)).

Synaptic Anatomy

Locations of synapses were derived using an algorithm described in the companion article (Reimann et al., 2015). The algorithm eliminates appositions that do not comply with the multi-synapse and plasticity reserve rules and ensures compatibility with observed biological bouton densities.

Synaptic Physiology

Excitatory synaptic transmission was modeled using both AMPA and NMDA receptor kinetics (Fuhrmann et al., 2002; Häusser and Roth, 1997; Markram et al., 1998; Ramaswamy et al., 2012; Tsodyks and Markram, 1997). Inhibitory synaptic transmission was modeled with a combination of GABA_A and GABA_B receptor kinetics (Gupta et al., 2000; Khazipov et al., 1995; De Koninck and Mody, 1997; Mott et al., 1999). Stochastic synaptic transmission was implemented as a two-state Markov model of dynamic synaptic release, a stochastic implementation of the Tsodyks-Markram dynamic synapse model (Fuhrmann et al., 2002; Tsodyks and Markram, 1997). Biological parameter ranges for the four model parameters were taken from experimental values for synaptic connections between specific m- and me-types or between larger categories of pre- and postsynaptic neurons (see Figure 9). Spontaneous miniature PSCs were modeled by implementing an independent Poisson process for each individual synapse that triggered release at rates (λ_{spont}) determined by the experimental data (Ling and Benardo, 1999; Simkus and Stricker, 2002).

Thalamic Innervation

Thalamic input was reconstructed using experimental data for ventro-posterior medial (VPM) axon bouton density profiles in rat barrel cortex (Meyer et al., 2010b), synapses per connection, and approximate numbers of incoming fibers. Synapse locations were determined using a variant of the connectome algorithm (Reimann et al., 2015; see [Supplemental Experimental Procedures](#)). Synapses were assigned to incoming fibers based on a Gaussian probability centered around each fiber.

Simulation

Microcircuit Simulation

The reconstructed microcircuit was simulated using the NEURON simulation package, augmented for execution on the supercomputer (Hines and Carnevale, 1997; Hines et al., 2008a, 2008b), together with additional custom tools to handle the setup and configuration of the microcircuit and the output of results.

In Silico Experiments

Depolarization was achieved by simulating current injection at the neuron soma. Currents were expressed as percent of first spike threshold for each neuron. Changes in $[\text{Ca}^{2+}]_o$ were simulated by changing the use parameter of synaptic transmission according to three curves for specific m-types (see Figure S15). Neuronal in silico knockout experiments were performed by hyperpolarizing the target population with somatic current injection (-100% threshold). Thalamic fiber stimulations were performed on circular clusters of minicolumns. The methods used to replicate previous in vivo experiments are described in the [Supplemental Experimental Procedures](#).

Data Analysis

Anatomical and physiological data analysis were performed using a custom suite of Python-based tools operating on a Linux cluster ([Supplemental Experimental Procedures](#)). The same analysis as described in Kasthuri et al. (2015) was applied to compare results between a saturated EM reconstruction and the digital reconstruction that we generated (see Figure S20). PSPs were measured at the somata or dendrites of randomly selected pairs of neurons (30 trials). PSTHs were computed from all neurons in the circuit and were normalized by neuron number and time bin to express the average instantaneous firing rate. Mean spike-spike correlations were calculated as the histogram of intervals between all spike times of two different cells (bin size 1 ms). Evoked regenerative activity was defined as activity in which peak activity (PSTH) within 100 ms after stimulus of the outermost 20 minicolumns exceeded 30 Hz and 70% of the activity of the 20 central minicolumns in the 10 ms after stimulus. Spike rasters show spike events at the locations within the layers where they occurred (for clarity, only a fraction of spikes are plotted).

Supercomputing

Reconstruction and simulation workflows, such as neuron model optimization, circuit reconstruction, and network simulation, were executed on supercomputers. The systems used included an IBM Blue Gene/L (until 2009), a CADMOS 4-rack IBM Blue Gene/P (until 2013), a CADMOS 1-rack IBM Blue Gene/Q (until 2014), and the Blue Brain IV operated by the Swiss National Supercomputing Center (CSCS) on behalf of the Blue Brain Project, ranked the 100th most powerful supercomputing system (Top500, June 2015). Blue Brain IV includes a 4-rack IBM Blue Gene/Q, IBM Blue Gene Active Storage, and a 40-node Linux cluster for post-processing, analysis, and visualization, fully interconnected using Infiniband technology and a GPFS file system with 4.2 Petabyte raw storage (Schürmann et al., 2014).

Visualization

Large circuits and simulations in high resolution were visualized using a custom-developed tool, RTNeuron (Hernando et al., 2012). High-quality, static images of small neural circuits, individual neurons, and synaptic spines and boutons were created using Maya 3D animation software (Autodesk, San Rafael, California, USA).

Software Development

Data integration and post processing as well as reconstruction, simulation, analysis, and visualization of neuronal network models used >30 software applications, integrated into automated and semi-automated workflows. Development was supported by a comprehensive development environment based on best practices for version control (git), code review (gerrit), and continuous building, testing, packaging, and deployment (Jenkins).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 20 figures, six tables, and three movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.029>.

AUTHOR CONTRIBUTIONS

H.M. conceived and led the study. F.S., S.L.H., I.S., and J.D. co-led the study. H.M. planned and supervised experiments, data integration, strategies and algorithms, model building, in silico experiments, and analysis. F.S. planned and supervised the development of algorithms, software and workflows, computing infrastructure, and technical integration. E.M., S.R., M.W.R., and S.L.H. drove and co-supervised the integration of the data, tools, models,

simulations, and analyses. H.M., S.R., E.M., and M.W.R. wrote the manuscript. R.W. edited the manuscript. A detailed listing of author contributions is available in the [Supplemental Information](#).

ACKNOWLEDGMENTS

The work was supported by funding from the EPFL to the Laboratory of Neural Microcircuitry (LNMC) and funding from the ETH Domain for the Blue Brain Project (BBP). Additional support was provided by funding for the Human Brain Project from the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement no. 604102 (HBP). Further funding came from The Gatsby Charitable Foundation; the Cajal Blue Brain Project, Ministerio de Economía y Competitividad Spanish Ministry of Education and Science; and an EPFL-Hebrew University Collaborative Grant. In the years 2005–2009, the Blue Gene/L system was funded by the EPFL. Financial support for the subsequent CADMOS Blue Gene/P and Blue Gene/Q systems was provided by the Canton of Geneva, Canton of Vaud, Hans Wilsdorf Foundation, Louis-Jeantet Foundation, University of Geneva, University of Lausanne, and École Polytechnique Fédérale de Lausanne. The BlueBrain IV BlueGene/Q system is financed by ETH Board Funding to the Blue Brain Project as a National Research Infrastructure and hosted at the Swiss National Supercomputing Center (CSCS). A large proportion of the data used in this study was generated at the Weizmann Institute for Science, Israel between 1996 and 2002 through the support of Thomas McKenna from the Office of Naval Research, USA.

We further acknowledge: Georges Abou-Jaoudé for generating the first Blue Brain visualizations and inspirational discussions on visualization; Shadi Akiki for contributions to circuit building and analysis; Katia Antonello for reconstructions; Haroon Anwar for contributions to morphology repair; Ricardo Auhing for contributions to databasing; Christiane Debono for administrative support; Raphael Holzer for development of Igor programs; Luca Gambazzi for contributions to ion channel and morphology databasing; Sonia Garcia for experiment support and reconstructions; Marc-Oliver Gewaltig for feedback on the modeling; the late Philip Goodman for multiple contributions over the years, starting with the first Beowulf computer cluster for brain simulation and for his statistical analyses of gene expression; Rony Hatteland for contributions to morphology cloning; David Horrigan for helpful discussions; Asif Jan for contributions to databasing; James Kozloski for contributions to touch detection; Tara Mahfoud for helpful discussions; Thomas McColgan for contributions to morphology repair; Ruben Moor for contributions to feature extraction; Charles Peck for contributions to touch detection; Jose M. (Chema) Peña for contributions to segmentation; Sandrine Romand for data on the maturation of pyramidal cells; Niklas Schmücker for contributions to the cell building process; Mohit Srivastava for contributions to morphological reconstructions; Garik Suess for contributions to the simulation; Anirudh Vij for contributions to the analysis software; Barthelemy von Haller for contributions on software and workflows for analysis; Anna Traussnig for contributions to circuit analysis; Tahir Uddin for contributions to morphological reconstructions; Alex Thomson for her pioneering work on synaptically coupled neurons that inspired our work, for many years of discussion and feedback, for detailed comments on early versions of the manuscript, and for continuous encouragement; the many other colleagues at IBM for stimulating discussions over the years since 2005; Bert Sakmann for his invaluable guidance, critical feedback, and discussions at the start of this study in 1994; and Kamila Markram for helpful discussions, limitless encouragement, support, and patience.

Received: December 16, 2014

Revised: May 4, 2015

Accepted: September 11, 2015

Published: October 8, 2015

REFERENCES

Ali, A.B., Bannister, A.P., and Thomson, A.M. (2007). Robust correlations between action potential duration and the properties of synaptic connections in layer 4 interneurons in neocortical slices from juvenile rats and adult rat and cat. *J. Physiol.* 580, 149–169.

Amithai, Y. (2001). Thalamocortical synaptic connections: efficacy, modulation, inhibition and plasticity. *Rev. Neurosci.* 12, 159–173.

Amzica, F., Massimini, M., and Manfredi, A. (2002). Spatial buffering during slow and paroxysmal sleep oscillations in cortical networks of glial cells in vivo. *J. Neurosci.* 22, 1042–1053.

Angulo, M.C., Rossier, J., and Audinat, E. (1999). Postsynaptic glutamate receptors and integrative properties of fast-spiking interneurons in the rat neocortex. *J. Neurophysiol.* 82, 1295–1302.

Anwar, H., Riachi, I., Hill, S., Schurmann, F., and Markram, H. (2009). An approach to capturing neuron morphological diversity. In *Computational Modeling Methods for Neuroscientists* (The MIT Press), pp. 211–231.

Ascoli, G.A., Alonso-Nanclares, L., Anderson, S.A., Barrionuevo, G., Benavides-Piccone, R., Burkhalter, A., Buzsáki, G., Cauli, B., Defelipe, J., Fairén, A., et al.; Petilla Interneuron Nomenclature Group (2008). Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* 9, 557–568.

Bannister, A.P., and Thomson, A.M. (2007). Dynamic properties of excitatory synaptic connections involving layer 4 pyramidal cells in adult rat and cat neocortex. *Cereb. Cortex* 17, 2190–2203.

Beaulieu, C. (1993). Numerical data on neocortical neurons in adult rat, with special reference to the GABA population. *Brain Res.* 609, 284–292.

Beaulieu, C., and Colonnier, M. (1983). The number of neurons in the different laminae of the binocular and monocular regions of area 17 in the cat, Canada. *J. Comp. Neurol.* 217, 337–344.

Beggs, J.M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci.* 23, 11167–11177.

Beierlein, M., and Connors, B.W. (2002). Short-term dynamics of thalamocortical and intracortical synapses onto layer 6 neurons in neocortex. *J. Neurophysiol.* 88, 1924–1932.

Beierlein, M., Gibson, J.R., and Connors, B.W. (2000). A network of electrically coupled interneurons drives synchronized inhibition in neocortex. *Nat. Neurosci.* 3, 904–910.

Beierlein, M., Fall, C.P., Rinzel, J., and Yuste, R. (2002). Thalamocortical bursts trigger recurrent activity in neocortical networks: layer 4 as a frequency-dependent gate. *J. Neurosci.* 22, 9885–9894.

Beierlein, M., Gibson, J.R., and Connors, B.W. (2003). Two dynamically distinct inhibitory networks in layer 4 of the neocortex. *J. Neurophysiol.* 90, 2987–3000.

Berger, T., Larkum, M.E., and Lüscher, H.-R. (2001). High I(h) channel density in the distal apical dendrite of layer V pyramidal cells increases bidirectional attenuation of EPSPs. *J. Neurophysiol.* 85, 855–868.

Blatow, M., Rozov, A., Katona, I., Hormuzdi, S.G., Meyer, A.H., Whittington, M.A., Caputi, A., and Monyer, H. (2003). A novel network of multipolar bursting interneurons generates theta frequency oscillations in neocortex. *Neuron* 38, 805–817.

Borst, J.G.G. (2010). The low synaptic release probability in vivo. *Trends Neurosci.* 33, 259–266.

Boyd, J., and Matsubara, J. (1991). Intrinsic connections in cat visual cortex: a combined anterograde and retrograde tracing study. *Brain Res.* 560, 207–215.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208.

Bruno, R.M., and Sakmann, B. (2006). Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* 312, 1622–1627.

Callaway, E.M. (2008). Transneuronal circuit tracing with neurotropic viruses. *Curr. Opin. Neurobiol.* 18, 617–623.

Carnevale, N.T., and Hines, M.L. (2006). *The NEURON Book* (New York, NY, USA: Cambridge University Press).

Cauli, B., Audinat, E., Lambolez, B., Angulo, M.C., Ropert, N., Tsuzuki, K., Hestrin, S., and Rossier, J. (1997). Molecular and physiological diversity of cortical nonpyramidal cells. *J. Neurosci.* 17, 3894–3906.

- Cauli, B., Porter, J.T., Tsuzuki, K., Lambolez, B., Rossier, J., Quenet, B., and Audinat, E. (2000). Classification of fusiform neocortical interneurons based on unsupervised clustering. *Proc. Natl. Acad. Sci. USA* 97, 6144–6149.
- Celikel, T., Szostak, V.A., and Feldman, D.E. (2004). Modulation of spike timing by sensory deprivation during induction of cortical map plasticity. *Nat. Neurosci.* 7, 534–541.
- Celio, M.R. (1986). Parvalbumin in most gamma-aminobutyric acid-containing neurons of the rat cerebral cortex. *Science* 231, 995–997.
- Chen, N., Sugihara, H., and Sur, M. (2015). An acetylcholine-activated micro-circuit drives temporal dynamics of cortical activity. *Nat. Neurosci.* 18, 892–902.
- Cobb, S.R., Halasy, K., Vida, I., Nyíri, G., Tamás, G., Buhl, E.H., and Somogyi, P. (1997). Synaptic effects of identified interneurons innervating both interneurons and pyramidal cells in the rat hippocampus. *Neuroscience* 79, 629–648.
- Condé, F., Lund, J.S., Jacobowitz, D.M., Baimbridge, K.G., and Lewis, D.A. (1994). Local circuit neurons immunoreactive for calretinin, calbindin D-28k or parvalbumin in monkey prefrontal cortex: distribution and morphology. *J. Comp. Neurol.* 341, 95–116.
- Connors, B.W., and Gutnick, M.J. (1990). Intrinsic firing patterns of diverse neocortical neurons. *Trends Neurosci.* 13, 99–104.
- Constantinople, C.M., and Bruno, R.M. (2011). Effects and mechanisms of wakefulness on local cortical networks. *Neuron* 69, 1061–1068.
- Constantinople, C.M., and Bruno, R.M. (2013). Deep cortical layers are activated directly by thalamus. *Science* 340, 1591–1594.
- Cragg, B.G. (1967). The density of synapses and neurones in the motor and visual areas of the cerebral cortex. *J. Anat.* 101, 639–654.
- Crick, F.H. (1979). Thinking about the brain. *Sci. Am.* 241, 219–232.
- Cunningham, M.O., Whittington, M.A., Bibbig, A., Roopun, A., LeBeau, F.E.N., Vogt, A., Monyer, H., Buhl, E.H., and Traub, R.D. (2004). A role for fast rhythmic bursting neurons in cortical gamma oscillations in vitro. *Proc. Natl. Acad. Sci. USA* 101, 7152–7157.
- Datwani, A., Iwasato, T., Itoharu, S., and Erzurumlu, R.S. (2002). NMDA receptor-dependent pattern transfer from afferents to postsynaptic cells and dendritic differentiation in the barrel cortex. *Mol. Cell. Neurosci.* 21, 477–492.
- De Koninck, Y., and Mody, I. (1997). Endogenous GABA activates small-conductance K⁺ channels underlying slow IPSCs in rat hippocampal neurons. *J. Neurophysiol.* 77, 2202–2208.
- DeFelipe, J. (1993). Neocortical neuronal diversity: chemical heterogeneity revealed by colocalization studies of classic neurotransmitters, neuropeptides, calcium-binding proteins, and cell surface molecules. *Cereb. Cortex* 3, 273–289.
- DeFelipe, J., and Jones, E.G. (1988). *Cajal on the Cerebral Cortex: An Annotated Translation of the Complete Writings* (Oxford University Press).
- DeFelipe, J., Alonso-Nanclares, L., and Arellano, J.I. (2002). Microstructure of the neocortex: comparative aspects. *J. Neurocytol.* 31, 299–316.
- DeFelipe, J., López-Cruz, P.L., Benavides-Piccione, R., Bielza, C., Larrañaga, P., Anderson, S., Burkhalter, A., Cauli, B., Fairén, A., Feldmeyer, D., et al. (2013). New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* 14, 202–216.
- Delattre, V., Keller, D., Perich, M., Markram, H., and Müller, E.B. (2015). Network-timing-dependent plasticity. *Front. Cell. Neurosci.* 9, 220.
- Denk, W., and Horstmann, H. (2004). Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* 2, e329.
- Douglas, R.J., and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451.
- Druckmann, S., Banitt, Y., Gidon, A., Schürmann, F., Markram, H., and Segev, I. (2007). A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Front. Neurosci.* 1, 7–18.
- Druckmann, S., Berger, T.K., Schürmann, F., Hill, S., Markram, H., and Segev, I. (2011). Effective stimuli for constructing reliable neuron models. *PLoS Comput. Biol.* 7, e1002133.
- Druckmann, S., Hill, S., Schürmann, F., Markram, H., and Segev, I. (2013). A hierarchical structure of cortical interneuron electrical diversity revealed by automated statistical analysis. *Cereb. Cortex* 23, 2994–3006.
- Dumitriu, D., Cossart, R., Huang, J., and Yuste, R. (2007). Correlation between axonal morphologies and synaptic input kinetics of interneurons from mouse visual cortex. *Cereb. Cortex* 17, 81–91.
- Fairén, A., DeFelipe, J., and Regidor, J. (1984). Nonpyramidal neurons: general account. In *Cerebral Cortex, Volume 1, Cellular Components of the Cerebral Cortex*, A. Peters and E.G. Jones, eds. (Plenum Press), pp. 201–253.
- Feldmeyer, D., Egger, V., Lübke, J., and Sakmann, B. (1999). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *J. Physiol.* 521, 169–190.
- Feldmeyer, D., Lübke, J., Silver, R.A., and Sakmann, B. (2002). Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column. *J. Physiol.* 538, 803–822.
- Feldmeyer, D., Lübke, J., and Sakmann, B. (2006). Efficacy and connectivity of intracolumnar pairs of layer 2/3 pyramidal cells in the barrel cortex of juvenile rats. *J. Physiol.* 575, 583–602.
- Frick, A., Feldmeyer, D., and Sakmann, B. (2007). Postnatal development of synaptic transmission in local networks of L5A pyramidal neurons in rat somatosensory cortex. *J. Physiol.* 585, 103–116.
- Frick, A., Feldmeyer, D., Helmstaedter, M., and Sakmann, B. (2008). Monosynaptic connections between pairs of L5A pyramidal neurons in columns of juvenile rat somatosensory cortex. *Cereb. Cortex* 18, 397–406.
- Fuhrmann, G., Segev, I., Markram, H., and Tsodyks, M. (2002). Coding of temporal information by activity-dependent synapses. *J. Neurophysiol.* 87, 140–148.
- Gentet, L.J., Avermann, M., Matyas, F., Staiger, J.F., and Petersen, C.C.H. (2010). Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* 65, 422–435.
- Gentet, L.J., Kremer, Y., Taniguchi, H., Huang, Z.J., Staiger, J.F., and Petersen, C.C.H. (2012). Unique functional properties of somatostatin-expressing GABAergic neurons in mouse barrel cortex. *Nat. Neurosci.* 15, 607–612.
- Ghosh, A., and Shatz, C.J. (1993). A role for subplate neurons in the patterning of connections from thalamus to neocortex. *Development* 117, 1031–1047.
- Gil, Z., Connors, B.W., and Amitai, Y. (1999). Efficacy of thalamocortical and intracortical synaptic connections: quanta, innervation, and reliability. *Neuron* 23, 385–397.
- Glenn, L.L., Hada, J., Roy, J.P., Deschênes, M., and Steriade, M. (1982). Anterograde tracer and field potential analysis of the neocortical layer I projection from nucleus ventralis medialis of the thalamus in cat. *Neuroscience* 7, 1861–1877.
- Gonchar, Y., and Burkhalter, A. (1997). Three distinct families of GABAergic neurons in rat visual cortex. *Cereb. Cortex* 7, 347–358.
- Gonchar, Y., Wang, Q., Burkhalter, A., Gonchar, Y., Wang, Q., and Burkhalter, A. (2007). Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Front. Neuroanat.* 1, 3.
- Grange, P., Bohland, J.W., Okaty, B.W., Sugino, K., Bokil, H., Nelson, S.B., Ng, L., Hawrylycz, M., and Mitra, P.P. (2014). Cell-type-based model explaining coexpression patterns of genes in the brain. *Proc. Natl. Acad. Sci. USA* 111, 5397–5402.
- Gupta, A., Wang, Y., and Markram, H. (2000). Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* 287, 273–278.
- Haider, B., Duque, A., Hasenstaub, A.R., and McCormick, D.A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* 26, 4535–4545.

- Hallman, L.E., Schofield, B.R., and Lin, C.-S. (1988). Dendritic morphology and axon collaterals of corticotectal, corticopontine, and callosal neurons in layer V of primary visual cortex of the hooded rat. *J. Comp. Neurol.* 272, 149–160.
- Harris, K.D., and Shepherd, G.M.G. (2015). The neocortical circuit: themes and variations. *Nat. Neurosci.* 18, 170–181.
- Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., and McCormick, D.A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* 47, 423–435.
- Häusser, M., and Roth, A. (1997). Estimating the time course of the excitatory synaptic conductance in neocortical pyramidal cells using a novel voltage jump method. *J. Neurosci.* 17, 7606–7625.
- Hay, E., Hill, S., Schürmann, F., Markram, H., and Segev, I. (2011). Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Comput. Biol.* 7, e1002107.
- Heinemann, U., Lux, H.D., and Gutnick, M.J. (1977). Extracellular free calcium and potassium during paroxysmal activity in the cerebral cortex of the cat. *Exp. Brain Res.* 27, 237–243.
- Hendry, S.H., Jones, E.G., and Emson, P.C. (1984). Morphology, distribution, and synaptic relations of somatostatin- and neuropeptide Y-immunoreactive neurons in rat and monkey neocortex. *J. Neurosci.* 4, 2497–2517.
- Hendry, S.H.C., Jones, E.G., Emson, P.C., Lawson, D.E.M., Heizmann, C.W., and Streit, P. (1989). Two classes of cortical GABA neurons defined by differential calcium binding protein immunoreactivities. *Exp. Brain Res.* 76, 467–472.
- Hernando, J., Schürmann, F., and Pastor, L. (2012). Towards real-time visualization of detailed neural tissue models: View frustum culling for parallel rendering. *BioVis*, 25–32.
- Hestrin, S., and Armstrong, W.E. (1996). Morphology and physiology of cortical neurons in layer I. *J. Neurosci.* 16, 5290–5300.
- Hevner, R.F., Shi, L., Justice, N., Hsueh, Y., Sheng, M., Smiga, S., Bulfone, A., Goffinet, A.M., Campagnoni, A.T., and Rubenstein, J.L.R. (2001). *Tbr1* regulates differentiation of the preplate and layer 6. *Neuron* 29, 353–366.
- Hill, S.L., Wang, Y., Riachi, I., Schürmann, F., and Markram, H. (2012). Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proc. Natl. Acad. Sci. USA* 109, E2885–E2894.
- Hines, M.L., and Carnevale, N.T. (1997). The NEURON simulation environment. *Neural Comput.* 9, 1179–1209.
- Hines, M.L., Markram, H., and Schürmann, F. (2008a). Fully implicit parallel simulation of single neurons. *J. Comput. Neurosci.* 25, 439–448.
- Hines, M.L., Eichner, H., and Schürmann, F. (2008b). Neuron splitting in compute-bound parallel network simulations enables runtime scaling with twice as many processors. *J. Comput. Neurosci.* 25, 203–210.
- Hines, M., Kumar, S., and Schürmann, F. (2011). Comparison of neuronal spike exchange methods on a Blue Gene/P supercomputer. *Front. Comput. Neurosci.* 5, 49.
- Holmgren, C., Harkany, T., Svennenfors, B., and Zilberter, Y. (2003). Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol.* 551, 139–153.
- Honey, C.J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. USA* 104, 10240–10245.
- Horton, J.C., and Adams, D.L. (2005). The cortical column: a structure without a function. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360, 837–862.
- Hughes, S.W., Cope, D.W., Blethyn, K.L., and Crunelli, V. (2002). Cellular mechanisms of the slow (<1 Hz) oscillation in thalamocortical neurons in vitro. *Neuron* 33, 947–958.
- Jones, H.C., and Keep, R.F. (1988). Brain fluid calcium concentration and response to acute hypercalcaemia during development in the rat. *J. Physiol.* 402, 579–593.
- Kapfer, C., Glickfeld, L.L., Atallah, B.V., and Scanziani, M. (2007). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nat. Neurosci.* 10, 743–753.
- Karagiannis, A., Gallopin, T., Dávid, C., Battaglia, D., Geoffroy, H., Rossier, J., Hillman, E.M.C., Staiger, J.F., and Cauli, B. (2009). Classification of NPY-expressing neocortical interneurons. *J. Neurosci.* 29, 3642–3659.
- Karube, F., Kubota, Y., and Kawaguchi, Y. (2004). Axon branching and synaptic bouton phenotypes in GABAergic nonpyramidal cell subtypes. *J. Neurosci.* 24, 2853–2865.
- Kasper, E.M., Larkman, A.U., Lübke, J., and Blakemore, C. (1994). Pyramidal neurons in layer 5 of the rat visual cortex. II. Development of electrophysiological properties. *J. Comp. Neurol.* 339, 475–494.
- Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al. (2015). Saturated Reconstruction of a Volume of Neocortex. *Cell* 162, 648–661.
- Kawaguchi, Y., and Kondo, S. (2002). Parvalbumin, somatostatin and cholecystokinin as chemical markers for specific GABAergic interneuron types in the rat frontal cortex. *J. Neurocytol.* 31, 277–287.
- Kawaguchi, Y., and Kubota, Y. (1993). Correlation of physiological subgroupings of nonpyramidal cells with parvalbumin- and calbindinD28k-immunoreactive neurons in layer V of rat frontal cortex. *J. Neurophysiol.* 70, 387–396.
- Kawaguchi, Y., and Kubota, Y. (1997). GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* 7, 476–486.
- Kawaguchi, Y., Karube, F., and Kubota, Y. (2006). Dendritic branch typing and spine expression patterns in cortical nonpyramidal cells. *Cereb. Cortex* 16, 696–711.
- Keller, A., and Carlson, G.C. (1999). Neonatal whisker clipping alters intracortical, but not thalamocortical projections, in rat barrel cortex. *J. Comp. Neurol.* 412, 83–94.
- Khazen, G., Hill, S.L., Schürmann, F., and Markram, H. (2012). Combinatorial expression rules of ion channel genes in juvenile rat (*Rattus norvegicus*) neocortical neurons. *PLoS ONE* 7, e34786.
- Khazipov, R., Congar, P., and Ben-Ari, Y. (1995). Hippocampal CA1 lacunosum-moleculare interneurons: modulation of monosynaptic GABAergic IPSCs by presynaptic GABAB receptors. *J. Neurophysiol.* 74, 2126–2137.
- Killackey, H.P., Gould, H.J., 3rd, Cusick, C.G., Pons, T.P., and Kaas, J.H. (1983). The relation of corpus callosum connections to architectonic fields and body surface maps in sensorimotor cortex of new and old world monkeys. *J. Comp. Neurol.* 219, 384–419.
- Kisvárdy, Z.F., Martin, K.A., Whitteridge, D., and Somogyi, P. (1985). Synaptic connections of intracellularly filled clutch cells: a type of small basket cell in the visual cortex of the cat. *J. Comp. Neurol.* 241, 111–137.
- Klausberger, T., Magill, P.J., Márton, L.F., Roberts, J.D.B., Cobden, P.M., Buzsáki, G., and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. *Nature* 421, 844–848.
- Le Bé, J.-V., Silberberg, G., Wang, Y., and Markram, H. (2007). Morphological, electrophysiological, and synaptic properties of corticocortical pyramidal cells in the neonatal rat neocortex. *Cereb. Cortex* 17, 2204–2213.
- Krimer, L.S., Zaitsev, A.V., Czanner, G., Kröner, S., González-Burgos, G., Poyshcheva, N.V., Iyengar, S., Barriónuevo, G., and Lewis, D.A. (2005). Cluster analysis-based physiological classification and morphological properties of inhibitory neurons in layers 2–3 of monkey dorsolateral prefrontal cortex. *J. Neurophysiol.* 94, 3009–3022.
- Larkman, A.U. (1991a). Dendritic morphology of pyramidal neurones of the visual cortex of the rat: I. Branching patterns. *J. Comp. Neurol.* 306, 307–319.
- Larkman, A.U. (1991b). Dendritic morphology of pyramidal neurones of the visual cortex of the rat: III. Spine distributions. *J. Comp. Neurol.* 306, 332–343.
- Larkum, M.E., Zhu, J.J., and Sakmann, B. (2001). Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. *J. Physiol.* 533, 447–466.

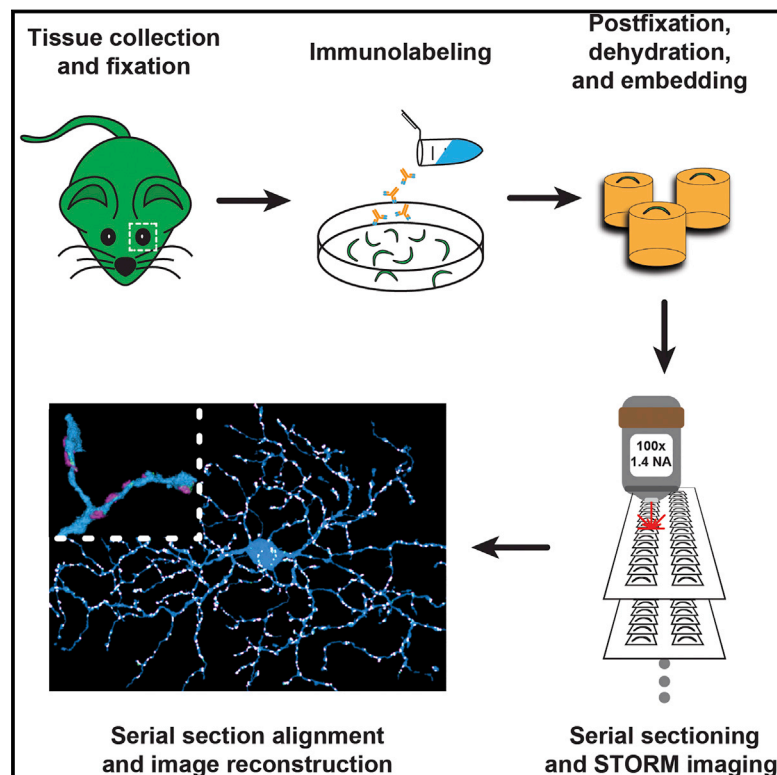
- Lee, S.-H., and Dan, Y. (2012). Neuromodulation of brain states. *Neuron* 76, 209–222.
- Lefort, S., Tómm, C., Floyd Sarria, J.-C., and Petersen, C.C.H. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron* 61, 301–316.
- Leinekugel, X., Khazipov, R., Cannon, R., Hirase, H., Ben-Ari, Y., and Buzsáki, G. (2002). Correlated bursts of activity in the neonatal hippocampus in vivo. *Science* 296, 2049–2052.
- Ling, D.S., and Benardo, L.S. (1999). Restrictions on inhibitory circuits contribute to limited recruitment of fast inhibition in rat neocortical pyramidal cells. *J. Neurophysiol.* 82, 1793–1807.
- Livingstone, M.S. (1996). Oscillatory firing and interneuronal correlations in squirrel monkey striate cortex. *J. Neurophysiol.* 75, 2467–2485.
- Lőrincz, M.L., Gunner, D., Bao, Y., Connelly, W.M., Isaac, J.T.R., Hughes, S.W., and Crunelli, V. (2015). A distinct class of slow (~0.2–2 Hz) intrinsically bursting layer 5 pyramidal neurons determines UP/DOWN state dynamics in the neocortex. *J. Neurosci.* 35, 5442–5458.
- Luczak, A., Barthó, P., Marguet, S.L., Buzsáki, G., and Harris, K.D. (2007). Sequential structure of neocortical spontaneous activity in vivo. *Proc. Natl. Acad. Sci. USA* 104, 347–352.
- Maffei, A., Nelson, S.B., and Turrigiano, G.G. (2004). Selective reconfiguration of layer 4 visual cortical circuitry by visual deprivation. *Nat. Neurosci.* 7, 1353–1359.
- Mao, B.-Q., Hamzei-Sichani, F., Aronov, D., Froemke, R.C., and Yuste, R. (2001). Dynamics of spontaneous activity in neocortical slices. *Neuron* 32, 883–898.
- Markram, H. (2008). Fixing the location and dimensions of functional neocortical columns. *HFSP J.* 2, 132–135.
- Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500, 409–440.
- Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. USA* 95, 5323–5328.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* 5, 793–807.
- Mason, A., Nicoll, A., and Stratford, K. (1991). Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *J. Neurosci.* 11, 72–84.
- Massimini, M., and Amzica, F. (2001). Extracellular calcium fluctuations and intracellular potentials in the cortex during the slow sleep oscillation. *J. Neurophysiol.* 85, 1346–1350.
- Mazurek, M.E., and Shadlen, M.N. (2002). Limits to the temporal fidelity of cortical spike rate signals. *Nat. Neurosci.* 5, 463–471.
- McCormick, D.A., Connors, B.W., Lighthall, J.W., and Prince, D.A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* 54, 782–806.
- McCormick, D.A., Shu, Y., Hasenstaub, A., Sanchez-Vives, M., Badoual, M., and Bal, T. (2003). Persistent cortical activity: mechanisms of generation and effects on neuronal excitability. *Cereb. Cortex* 13, 1219–1231.
- McGarry, L.M., Packer, A.M., Fino, E., Nikolenko, V., Sipky, T., and Yuste, R. (2010). Quantitative classification of somatostatin-positive neocortical interneurons identifies three interneuron subtypes. *Front. Neural Circuits* 4, 12.
- Mercer, A., West, D.C., Morris, O.T., Kirchhecker, S., Kerkhoff, J.E., and Thomson, A.M. (2005). Excitatory connections made by presynaptic cortico-cortical pyramidal cells in layer 6 of the neocortex. *Cereb. Cortex* 15, 1485–1496.
- Meyer, A.H., Katona, I., Blatow, M., Rozov, A., and Monyer, H. (2002). In vivo labeling of parvalbumin-positive interneurons and analysis of electrical coupling in identified neurons. *J. Neurosci.* 22, 7055–7064.
- Meyer, H.S., Wimmer, V.C., Oberlaender, M., de Kock, C.P.J., Sakmann, B., and Helmstaedter, M. (2010a). Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissa cortex. *Cereb. Cortex* 20, 2277–2286.
- Meyer, H.S., Wimmer, V.C., Hemberger, M., Bruno, R.M., de Kock, C.P., Frick, A., Sakmann, B., and Helmstaedter, M. (2010b). Cell type-specific thalamic innervation in a column of rat vibrissa cortex. *Cereb. Cortex* 20, 2287–2303.
- Meyer, H.S., Schwarz, D., Wimmer, V.C., Schmitt, A.C., Kerr, J.N.D., Sakmann, B., and Helmstaedter, M. (2011). Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5A. *Proc. Natl. Acad. Sci. USA* 108, 16807–16812.
- Micheva, K.D., and Smith, S.J. (2007). Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron* 55, 25–36.
- Micheva, K.D., Busse, B., Weiler, N.C., O'Rourke, N., and Smith, S.J. (2010). Single-synapse analysis of a diverse synapse population: proteomic imaging methods and markers. *Neuron* 68, 639–653.
- Migliore, M., Cannia, C., Lytton, W.W., Markram, H., and Hines, M.L. (2006). Parallel network simulations with NEURON. *J. Comput. Neurosci.* 21, 119–129.
- Mott, D.D., Li, Q., Okazaki, M.M., Turner, D.A., and Lewis, D.V. (1999). GABA_B-receptor-mediated currents in interneurons of the dentate-hilus border. *J. Neurophysiol.* 82, 1438–1450.
- Mountcastle, V.B. (1998). *Perceptual neuroscience: The cerebral cortex* (Harvard University Press).
- Nelson, S. (2002). Cortical microcircuits: diverse or canonical? *Neuron* 36, 19–27.
- Nevian, T., Larkum, M.E., Polsky, A., and Schiller, J. (2007). Properties of basal dendrites of layer 5 pyramidal neurons: a direct patch-clamp recording study. *Nat. Neurosci.* 10, 206–214.
- O'Connor, D.H., Huber, D., and Svoboda, K. (2009). Reverse engineering the mouse brain. *Nature* 461, 923–929.
- Oberlaender, M., de Kock, C.P.J., Bruno, R.M., Ramirez, A., Meyer, H.S., Derksen, V.J., Helmstaedter, M., and Sakmann, B. (2012). Cell type-specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissa cortex. *Cereb. Cortex* 22, 2375–2391.
- Ohana, O., and Sakmann, B. (1998). Transmitter release modulation in nerve terminals of rat neocortical pyramidal cells by intracellular calcium buffers. *J. Physiol.* 513, 135–148.
- Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.* 11, 535–537.
- Okun, M., Steinmetz, N.A., Cossell, L., Iacaruso, M.F., Ko, H., Barthó, P., Moore, T., Hofer, S.B., Mrcic-Flogel, T.D., Carandini, M., and Harris, K.D. (2015). Diverse coupling of neurons to populations in sensory cortex. *Nature* 521, 511–515.
- Oláh, S., Komlósi, G., Szabadics, J., Varga, C., Tóth, E., Barzó, P., and Tamás, G. (2007). Output of neurogliaform cells to various neuron types in the human and rat cerebral cortex. *Front. Neural Circuits* 1, 4.
- Packer, A.M., and Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J. Neurosci.* 31, 13260–13271.
- Perin, R., Berger, T.K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci. USA* 108, 5419–5424.
- Perrenoud, Q., Rossier, J., Geoffroy, H., Vitalis, T., and Gallopin, T. (2013). Diversity of GABAergic interneurons in layer VIa and VIb of mouse barrel cortex. *Cereb. Cortex* 23, 423–441.
- Peters, A. (1987). Number of Neurons and Synapses in Primary Visual Cortex. In *Cerebral Cortex*, E.G. Jones and A. Peters, eds. (Springer), pp. 267–294.
- Peters, A., and Kaiserman-Abramof, I.R. (1970). The small pyramidal neuron of the rat cerebral cortex. The perikaryon, dendrites and spines. *Am. J. Anat.* 127, 321–355.

- Ramaswamy, S., and Markram, H. (2015). Anatomy and physiology of the thick-tufted layer 5 pyramidal neuron. *Front. Cell. Neurosci.* 9, 233.
- Ramaswamy, S., Hill, S.L., King, J.G., Schürmann, F., Wang, Y., and Markram, H. (2012). Intrinsic morphological diversity of thick-tufted layer 5 pyramidal neurons ensures robust and invariant properties of in silico synaptic connections. *J. Physiol.* 590, 737–752.
- Ramaswamy, S., Courcol, J.-D., Abdellah, M., Adaszewski, S., Antille, N., Arsever, S., Guy Antoine, A.K., Bilgili, A., Brukau, Y., Chalimourda, A., et al. (2015). The neocortical Microcircuit collaboration portal: A resource for rat somatosensory cortex. *Front. Neural Circuits* 9, 44.
- Ramón y Cajal, S. (1909, 1911). *Histologie du Systeme Nerveux de l'Homme et des Vertébrés*. L. Azoulay, trans. Maloine, Paris.
- Reimann, M.W., Müller, E.B., Ramaswamy, S., and Markram, H. (2015). An algorithm to predict the connectome of neural microcircuits. *Front. Comput. Neurosci.* 9, 28.
- Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590.
- Reyes, A., and Sakmann, B. (1999). Developmental switch in the short-term modification of unitary EPSPs evoked in layer 2/3 and layer 5 pyramidal neurons of rat neocortex. *J. Neurosci.* 19, 3827–3835.
- Reyes, A., Lujan, R., Rozov, A., Burnashev, N., Somogyi, P., and Sakmann, B. (1998). Target-cell-specific facilitation and depression in neocortical circuits. *Nat. Neurosci.* 1, 279–285.
- Reyes-Puerta, V., Sun, J.-J., Kim, S., Kilb, W., and Luhmann, H.J. (2015). Laminar and Columnar Structure of Sensory-Evoked Multineuronal Spike Sequences in Adult Rat Barrel Cortex In Vivo. *Cereb. Cortex* 25, 2001–2021.
- Rinaldi, T., Silberberg, G., and Markram, H. (2008). Hyperconnectivity of local neocortical microcircuitry induced by prenatal exposure to valproic acid. *Cereb. Cortex* 18, 763–770.
- Romand, S., Wang, Y., Toledo-Rodríguez, M., and Markram, H. (2011). Morphological development of thick-tufted layer V pyramidal cells in the rat somatosensory cortex. *Front. Neuroanat.* 5, 5.
- Rozov, A., Burnashev, N., Sakmann, B., and Neher, E. (2001). Transmitter release modulation by intracellular Ca^{2+} buffers in facilitating and depressing nerve terminals of pyramidal cells in layer 2/3 of the rat neocortex indicates a target cell-specific difference in presynaptic calcium dynamics. *J. Physiol.* 531, 807–826.
- Rudy, B., Fishell, G., Lee, S., and Hjerling-Leffler, J. (2011). Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev. Neurobiol.* 71, 45–61.
- Sakata, S., and Harris, K.D. (2009). Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. *Neuron* 64, 404–418.
- Salinas, E., and Sejnowski, T.J. (2001). Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* 2, 539–550.
- Sancesario, G., Pisani, A., D'Angelo, V., Calabresi, P., and Bernardi, G. (1998). Morphological and functional study of dwarf neurons in the rat striatum. *Eur. J. Neurosci.* 10, 3575–3583.
- Sanchez-Vives, M.V., and McCormick, D.A. (2000). Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.* 3, 1027–1034.
- Santana, R., McGarry, L.M., Bielza, C., Larrañaga, P., and Yuste, R. (2013). Classification of neocortical interneurons using affinity propagation. *Front. Neural Circuits* 7, 185.
- Sarid, L., Bruno, R., Sakmann, B., Segev, I., and Feldmeyer, D. (2007). Modeling a layer 4-to-layer 2/3 module of a single column in rat neocortex: interweaving in vitro and in vivo experimental observations. *Proc. Natl. Acad. Sci. USA* 104, 16353–16358.
- Schroeder, C.E., and Foxe, J.J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–198.
- Schürmann, F., Delalondre, F., Kumbhar, P.S., Biddiscombe, J., Gila, M., Tacchella, D., Curioni, A., Metzler, B., Morjan, P., Fenkes, J., et al. (2014). Rebasin I/O for Scientific Computing: Leveraging Storage Class Memory in an IBM BlueGene/Q Supercomputer (Leipzig, Germany: Springer International Publishing Switzerland), pp. 331–347.
- Shu, Y., Hasenstaub, A., and McCormick, D.A. (2003). Turning on and off recurrent balanced cortical activity. *Nature* 423, 288–293.
- Sigalas, C., Rigas, P., Tsakanikas, P., and Skaliara, I. (2015). High-Affinity Nicotinic Receptors Modulate Spontaneous Cortical Up States In Vitro. *J. Neurosci.* 35, 11196–11208.
- Silberberg, G. (2008). Polysynaptic subcircuits in the neocortex: spatial and temporal diversity. *Curr. Opin. Neurobiol.* 18, 332–337.
- Silberberg, G., and Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron* 53, 735–746.
- Silberberg, G., Wu, C., and Markram, H. (2004). Synaptic dynamics control the timing of neuronal excitation in the activated neocortical microcircuit. *J. Physiol.* 556, 19–27.
- Silva, L.R., Amitai, Y., and Connors, B.W. (1991). Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science* 251, 432–435.
- Silver, R.A., Lubke, J., Sakmann, B., and Feldmeyer, D. (2003). High-probability unquantal transmission at excitatory synapses in barrel cortex. *Science* 302, 1981–1984.
- Simkus, C.R.L., and Stricker, C. (2002). Properties of mEPSCs recorded in layer II neurones of rat barrel cortex. *J. Physiol.* 545, 509–520.
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55, 349–374.
- Somogyi, P., Freund, T.F., and Cowey, A. (1982). The axo-axonic interneuron in the cerebral cortex of the rat, cat and monkey. *Neuroscience* 7, 2577–2607.
- Somogyi, P., Tamás, G., Lujan, R., and Buhl, E.H. (1998). Salient features of synaptic organisation in the cerebral cortex. *Brain Res. Brain Res. Rev.* 26, 113–135.
- Sporns, O., and Kötter, R. (2004). Motifs in brain networks. *PLoS Biol.* 2, e369.
- Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221.
- Stepanyants, A., Martinez, L.M., Ferecskó, A.S., and Kisvárdy, Z.F. (2009). The fractions of short- and long-range connections in the visual cortex. *Proc. Natl. Acad. Sci. USA* 106, 3555–3560.
- Steriade, M., Nuñez, A., and Amzica, F. (1993). A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.* 13, 3252–3265.
- Svoboda, K., Denk, W., Kleinfeld, D., and Tank, D.W. (1997). In vivo dendritic calcium dynamics in neocortical pyramidal neurons. *Nature* 385, 161–165.
- Szabadics, J., Varga, C., Molnár, G., Oláh, S., Barzó, P., and Tamás, G. (2006). Excitatory effect of GABAergic axo-axonic cells in cortical microcircuits. *Science* 311, 233–235.
- Tamás, G., Somogyi, P., and Buhl, E.H. (1998). Differentially interconnected networks of GABAergic interneurons in the visual cortex of the cat. *J. Neurosci.* 18, 4255–4270.
- Tan, A.Y.Y., Chen, Y., Scholl, B., Seidemann, E., and Priebe, N.J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature* 509, 226–229.
- Thomson, A.M., and Bannister, A.P. (1998). Postsynaptic pyramidal target selection by descending layer III pyramidal axons: dual intracellular recordings and biocytin filling in slices of rat neocortex. *Neuroscience* 84, 669–683.
- Thomson, A.M., and Deuchars, J. (1997). Synaptic interactions in neocortical local circuits: dual intracellular recordings in vitro. *Cereb. Cortex* 7, 510–522.
- Thomson, A.M., and Lamy, C. (2007). Functional maps of neocortical local circuitry. *Front. Neurosci.* 1, 19–42.
- Thomson, A.M., Deuchars, J., and West, D.C. (1993). Large, deep layer pyramidal-pyramidal single axon EPSPs in slices of rat motor cortex display paired pulse and frequency-dependent depression, mediated presynaptically

- and self-facilitation, mediated postsynaptically. *J. Neurophysiol.* 70, 2354–2369.
- Thomson, A.M., West, D.C., Hahn, J., and Deuchars, J. (1996). Single axon IPSPs elicited in pyramidal cells by three classes of interneurons in slices of rat neocortex. *J. Physiol.* 496, 81–102.
- Timofeev, I., Grenier, F., Bazhenov, M., Sejnowski, T.J., and Steriade, M. (2000). Origin of slow cortical oscillations in deafferented cortical slabs. *Cereb. Cortex* 10, 1185–1199.
- Toledo-Rodriguez, M., Blumenfeld, B., Wu, C., Luo, J., Attali, B., Goodman, P., and Markram, H. (2004). Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* 14, 1310–1327.
- Toledo-Rodriguez, M., Goodman, P., Illic, M., Wu, C., and Markram, H. (2005). Neuropeptide and calcium-binding protein gene expression profiles predict neuronal anatomical type in the juvenile rat. *J. Physiol.* 567, 401–413.
- Tsodyks, M.V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA* 94, 719–723.
- van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726.
- Wang, Z., and McCormick, D.A. (1993). Control of firing mode of corticotectal and corticopontine layer V burst-generating neurons by norepinephrine, acetylcholine, and 1S,3R-ACPD. *J. Neurosci.* 13, 2199–2216.
- Wang, Y., Gupta, A., Toledo-Rodriguez, M., Wu, C.Z., and Markram, H. (2002). Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cereb. Cortex* 12, 395–410.
- Wang, Y., Toledo-Rodriguez, M., Gupta, A., Wu, C., Silberberg, G., Luo, J., and Markram, H. (2004). Anatomical, physiological and molecular properties of Martinotti cells in the somatosensory cortex of the juvenile rat. *J. Physiol.* 561, 65–90.
- Wang, Y., Markram, H., Goodman, P.H., Berger, T.K., Ma, J., and Goldman-Rakic, P.S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* 9, 534–542.
- West, M.J., and Gundersen, H.J. (1990). Unbiased stereological estimation of the number of neurons in the human hippocampus. *J. Comp. Neurol.* 296, 1–22.
- Westerink, B.H.C., Hofsteede, H.M., Damsma, G., and de Vries, J.B. (1988). The significance of extracellular calcium for the release of dopamine, acetylcholine and amino acids in conscious rats, evaluated by brain microdialysis. *Naunyn Schmiedeberg's Arch. Pharmacol.* 337, 373–378.
- Wickersham, I.R., Finke, S., Conzelmann, K.-K., and Callaway, E.M. (2007). Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* 4, 47–49.
- Williams, S.R., and Mitchell, S.J. (2008). Direct measurement of somatic voltage clamp errors in central neurons. *Nat. Neurosci.* 11, 790–798.
- Williams, R.W., and Rakic, P. (1988). Three-dimensional counting: an accurate and direct method to estimate numbers of cells in sectioned material. *J. Comp. Neurol.* 278, 344–352.
- Wilson, N.R., Runyan, C.A., Wang, F.L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature* 488, 343–348.
- Wimmer, V.C., Bruno, R.M., de Kock, C.P.J., Kuner, T., and Sakmann, B. (2010). Dimensions of a projection column and architecture of VPM and POM axons in rat vibrissa cortex. *Cereb. Cortex* 20, 2265–2276.
- Yoshimura, Y., Kimura, F., and Tsumoto, T. (1999). Estimation of single channel conductance underlying synaptic transmission between pyramidal cells in the visual cortex. *Neuroscience* 88, 347–352.
- Yuste, R. (2005). Origin and classification of neocortical interneurons. *Neuron* 48, 524–527.
- Yuste, R., Tank, D.W., and Kleinfeld, D. (1997). Functional study of the rat cortical microcircuitry with voltage-sensitive dye imaging of neocortical slices. *Cereb. Cortex* 7, 546–558.
- Zagha, E., and McCormick, D.A. (2014). Neural control of brain state. *Curr. Opin. Neurobiol.* 29, 178–186.
- Zhu, J.J. (2000). Maturation of layer 5 neocortical pyramidal neurons: amplifying salient layer 1 and layer 4 inputs by Ca^{2+} action potentials in adult rat tuft dendrites. *J. Physiol.* 526, 571–587.

Mapping Synaptic Input Fields of Neurons with Super-Resolution Imaging

Graphical Abstract



Authors

Yaron M. Sigal, Colenso M. Speer,
Hazen P. Babcock, Xiaowei Zhuang

Correspondence

zhuang@chemistry.harvard.edu

In Brief

A super-resolution fluorescence-imaging platform for multi-color volumetric reconstruction of synapses and neurons in brain tissue is developed, enabling interrogation of neural circuitry at the nanoscale.

Highlights

- A super-resolution platform for volumetric tissue imaging and analysis is developed
- The platform allows synapses on neurons to be identified with molecular specificity
- On-Off DSGCs receive receptor subunit-specific GABAergic inputs
- On-Off DSGCs do not receive substantial monosynaptic glycinergic inputs

Mapping Synaptic Input Fields of Neurons with Super-Resolution Imaging

Yaron M. Sigal,^{1,4} Colenso M. Speer,^{1,4} Hazen P. Babcock,² and Xiaowei Zhuang^{1,2,3,*}

¹Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology

²Center for Brain Science

³Department of Physics

Harvard University, Cambridge, MA 02138, USA

⁴Co-first author

*Correspondence: zhuang@chemistry.harvard.edu

<http://dx.doi.org/10.1016/j.cell.2015.08.033>

SUMMARY

As a basic functional unit in neural circuits, each neuron integrates input signals from hundreds to thousands of synapses. Knowledge of the synaptic input fields of individual neurons, including the identity, strength, and location of each synapse, is essential for understanding how neurons compute. Here, we developed a volumetric super-resolution reconstruction platform for large-volume imaging and automated segmentation of neurons and synapses with molecular identity information. We used this platform to map inhibitory synaptic input fields of On-Off direction-selective ganglion cells (On-Off DSGCs), which are important for computing visual motion direction in the mouse retina. The reconstructions of On-Off DSGCs showed a GABAergic, receptor subtype-specific input field for generating direction selective responses without significant glycinergic inputs for mediating monosynaptic cross-over inhibition. These results demonstrate unique capabilities of this super-resolution platform for interrogating neural circuitry.

INTRODUCTION

Mapping synaptic connectivity at multiple scales, ranging from the synaptic fields of individual neurons to the wiring diagram of the whole brain, is important for understanding how neural circuits function and how circuit defects contribute to mental illness (Alivisatos et al., 2013; Morgan and Lichtman, 2013). An ideal platform for imaging synaptic connectivity should provide (1) high-resolution structural information for reliable identification of synaptic connections and accurate assignment of synapses to neurons; (2) the ability to image specific molecules, such as neurotransmitter receptors, important for determining synapse identity and properties; and (3) automated image segmentation capability for efficient analysis of large-volume reconstructions that capture entire neurons or circuits.

Both fluorescence microscopy and electron microscopy (EM) have been used for volumetric neural circuit reconstruction

(Helmstaedter, 2013; Kleinfeld et al., 2011; Lichtman and Denk, 2011). EM provides exquisite spatial resolution and membrane contrast for accurate synapse identification, and the high imaging speed of modern EM instruments allows increasingly larger volume reconstructions (Helmstaedter, 2013; Kleinfeld et al., 2011; Lichtman and Denk, 2011). However, because of the stringent fixation and sample preparation conditions required for high-quality EM imaging, labeling of endogenous synaptic proteins for determining the molecular identities and functional properties of synapses remains a difficult task for large-volume EM reconstructions. In addition, automated segmentation of EM images is still challenging and remains a bottleneck for scaling up neural circuit analysis, though substantial progress has been made on the development of automated EM image analysis and crowd-sourcing methods (Chklovskii et al., 2010; Helmstaedter, 2013; Jain et al., 2010). In comparison, fluorescence microscopy is compatible with immunohistochemistry and imaging of endogenous proteins over large volumes (Kleinfeld et al., 2011; Miyawaki, 2015), and multi-colored fluorescence signals can also help simplify the task of automated image segmentation for efficient data analysis. However, the diffraction-limited resolution of fluorescence microscopy can lead to substantial errors in the identification and assignment of synapses within reconstructed circuits.

Super-resolution fluorescence imaging overcomes the diffraction limit (Hell, 2007; Huang et al., 2010) and may enhance our ability to reconstruct neural circuits by integrating high image resolution for synapse identification and assignment, protein-specific labeling for determining the molecular properties of synapses, and multi-color imaging for efficient data analysis. Here, we developed a super-resolution reconstruction platform by combining stochastic optical reconstruction microscopy (STORM) (Huang et al., 2010; Rust et al., 2006) with serial ultra-thin sectioning for large-volume reconstruction of endogenous molecular targets in tissues and used this platform to image entire neurons and their synaptic inputs. We focused our studies on the inner plexiform layer (IPL) of the mouse retina where diverse classes of retinal ganglion cells (RGCs) integrate synaptic inputs (Anderson et al., 2011; Helmstaedter et al., 2013) to generate unique spatiotemporal representations of the visual scene (Gollisch and Meister, 2010). A classic example of such a computation is the determination of visual stimulus motion direction by On-Off direction-selective RGCs (On-Off DSGCs) (Vanev et al., 2012). The substantial prior knowledge of the structure

and function of this cell type allows validation of our method, while unresolved structural questions in this system provide an opportunity to test the ability of our approach to extract novel biological information. For example, On-Off DSGCs are known to receive asymmetric inhibitory GABAergic inputs from presynaptic starburst amacrine cells (SACs) during null-direction stimulus movement (Briggman et al., 2011; Fried et al., 2002; Wei et al., 2011), and the $\alpha 2$ subunit of GABA(A) receptor plays an important role in this direction selectivity (Auferkorte et al., 2012). In addition to GABAergic synapses, glycinergic signaling also impacts the response of On-Off DSGCs to the edges of moving stimuli (Caldwell et al., 1978; Jensen, 1999), likely reflecting crossover inhibition between the on and off sublaminae mediated by glycinergic amacrine cells (Kittila and Massey, 1995; Stasheff and Masland, 2002; Werblin, 2010). However, the structural basis of this crossover inhibition in On-Off DSGC circuits is incompletely understood, and it is unclear whether glycinergic interneurons make direct synaptic contacts onto On-Off DSGCs. To demonstrate the capabilities of our super-resolution platform, we reconstructed the inhibitory synaptic input fields of individual On-Off DSGCs and determined the spatial distribution and neurotransmitter receptor identity of the synapses therein. We also reconstructed the inhibitory input fields of two other types of retinal neurons, a small-field On-center RGC and a narrow-field amacrine cell, for comparative demonstration.

RESULTS

Volumetric, Multi-color Super-Resolution Reconstruction

We labeled neurons and synaptic proteins with spectrally distinct photoswitchable dyes for multi-color STORM imaging (Dempsey et al., 2011). For neuron labeling, we used mice expressing GFP or YFP in the cytoplasm of a sparse subset of retinal neurons (Feng et al., 2000) and labeled the dissected retinal tissue with anti-GFP antibodies. For marking inhibitory synapses, we used an antibody against an inhibitory synapse scaffolding protein, gephyrin, which anchors glycine and/or GABA receptors at postsynaptic terminals (Tyagarajan and Fritschy, 2014). For presynaptic counter-staining, we used a cocktail of antibodies against several active zone proteins, bassoon, piccolo, munc13-1, and ELKS, for dense labeling of all presynaptic terminals. Table S1 shows all of the antibodies tested in this work. We also included a general neuropil stain, wheat germ agglutinin (WGA), in a fourth color channel to produce images with dense information content to assist serial-section alignment.

For volumetric reconstruction, we embedded tissues in resin and used serial ultrathin sectioning, in combination with STORM imaging, to generate large-volume super-resolution images. Serial ultrathin sectioning not only facilitates large-volume fluorescence reconstruction of tissue samples but also allows the image resolution along the z direction (as defined by the section thickness) to be substantially higher than the diffraction limit, as has been demonstrated previously in array tomography and three-dimensional (3D) STED reconstructions (Micheva and Smith, 2007; Punge et al., 2008). The partial exposure of epitopes in samples embedded in acrylic resin also allows many

different synaptic proteins to be imaged through multiple rounds of post-embedding immunolabeling, which help identify synapses and characterize their molecular properties (Micheva et al., 2010; Micheva and Smith, 2007). However, the requirement of sample embedding for high-quality serial sectioning poses extra challenges for super-resolution imaging. Since STORM imaging relies on switching and localization of individual fluorophores to reconstruct super-resolution images (Huang et al., 2010; Rust et al., 2006), the resolution of a STORM image depends not only on the localization precision of individual fluorophores determined by their photon output but also on the localization density determined by the labeling density. Achieving optimal STORM resolution thus requires the labeling and embedding conditions to simultaneously retain optimal fluorophore properties and high-density labeling in resin-embedded samples. Resin embedding, however, substantially reduces the antigenicity of samples, which leads to a drastic reduction in antibody labeling density and severely compromises the image resolution achievable by STORM as we observed for tissues immunolabeled after acrylic resin embedding. Such low label densities, which are also evident in previous STORM images of tissue samples prepared using a similar post-embedding labeling approach (Nanguneri et al., 2012), prevent accurate tracing of neurons and identification of synapses using super-resolution imaging. We therefore explored pre-embedding immunofluorescence labeling (Punge et al., 2008) to increase the labeling density. We further tested various embedding materials and found that epoxy resin was excellent for maintaining the photon output of the fluorescent dyes. Finally, since optimal photoswitching of dye molecules requires access to a switching agent, such as thiol, we chemically etched the resin-embedded tissue sections using sodium ethoxide solution to expose the dyes to the thiol-containing imaging buffer.

Experimentally, we immunolabeled retinal tissues, performed an additional fixation step to crosslink the antibodies, dehydrated the samples, and embedded them in UltraBed epoxy resin (Figure 1A). The resin-embedded tissues were cut into 70 nm ultrathin sections, arrayed onto glass coverslips, and etched with sodium ethoxide (Figure 1A). Coverslips were imaged using a microscope setup that allowed automated imaging of entire arrays of sections, and both STORM and conventional images were collected for the same tissue sections. The xy-resolution of the STORM images was ~ 20 nm, and that of the conventional images was diffraction limited to ~ 200 – 300 nm, whereas the z resolution of both STORM and conventional images in this work was limited by the section thickness of 70 nm.

We developed an automated image analysis pipeline for processing STORM and conventional images, which included corrections of chromatic aberration and lens distortions using bead fiducials, as well as montage and serial-section alignment using scale-invariant feature transformation (SIFT) followed by elastic registration (Saalfeld et al., 2012) to generate large-volume reconstructions (Figure 1A) (see the [Experimental Procedures](#) for details). Volumetric STORM reconstructions of the IPL revealed efficient labeling throughout the sample, with neurons situated amidst hundreds of thousands of fluorescent clusters in each synaptic channel (Figures 1B, 1C, S1, and S2).

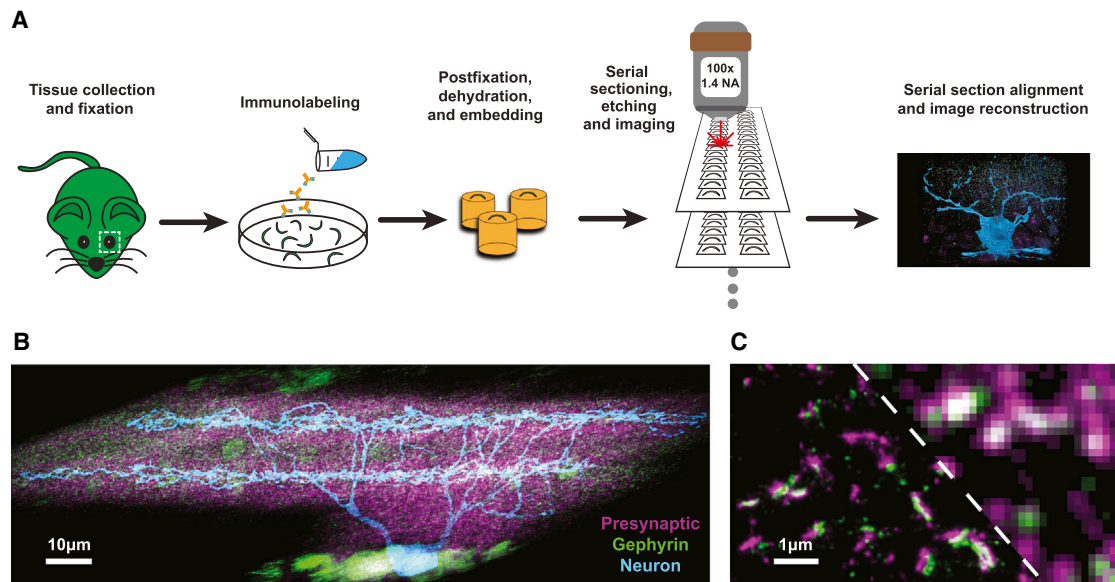


Figure 1. A Super-Resolution Imaging and Analysis Platform

(A) Tissues were dissected, fixed for immunohistochemical labeling, postfixed, dehydrated, and embedded in epoxy resin. Ultrathin sections were cut, arrayed on glass coverslips, and etched to expose fluorophores for STORM imaging. Individual serial sections were imaged and aligned to generate 3D reconstructions. (B) STORM maximum intensity projection of a volume ($2.3 \times 10^5 \mu\text{m}^3$) of the mouse IPL containing an On-Off DSGC (blue) amidst presynaptic (magenta) and gephyrin (green) clusters imaged using the platform.

(C) An enlarged image of synapses in a small region ($1 \mu\text{m}$ thickness) of the IPL. For comparison, the corresponding conventional image of the upper right portion (to the right of the dashed line) is presented.

See also [Figures S1](#) and [S2](#).

Synapse Identification

Taking advantage of the multi-color super-resolution fluorescent signals, we developed image segmentation algorithms for automated neuron and synapse identification and performed quantitative analysis of entire fields of molecularly identified synapses in our datasets. Labeled synaptic proteins appear as clusters of localizations in STORM images, but not all clusters in STORM images represent synapses ([Dani et al., 2010](#); [Specht et al., 2013](#)). For synapse identification, we measured the volume and signal density of all fluorescent clusters. In both presynaptic and postsynaptic (gephyrin) channels, these two parameters separated fluorescent clusters into two distinct populations ([Figure 2A](#)). We assigned the population of clusters with larger volumes as putative “synaptic” (S) structures and the other population with smaller volumes as putative “non-synaptic” (NS) structures. The vast majority ($\sim 91\%$) of the putative synaptic gephyrin clusters had closely apposed presynaptic clusters ([Figures S3A](#) and [S3C](#)), and example pairs of gephyrin and presynaptic clusters from this population clearly resembled synapses ([Figure 2B](#)), supporting our assignment. Of the putative non-synaptic population of gephyrin clusters, only a small fraction had a nearby presynaptic cluster ([Figure S3A](#)). Moreover, because some of these small gephyrin clusters were spatially close to the larger, paired gephyrin/presynaptic clusters and were thereby falsely identified as being paired, the automated pairing analysis of these small gephyrin clusters ([Figure S3A](#)) was less accurate than that for the larger synaptic clusters. Visual inspection showed that $\sim 90\%$ of these small gephyrin clusters were

unpaired and likely represent gephyrin-containing trafficking vesicles or background signals from non-specific antibody labeling, whereas the remaining small fraction of paired structures could represent small (potentially immature) synapses.

For the presynaptic clusters, even the synapse population contained a substantial fraction of clusters ($\sim 70\%$) that were not paired with gephyrin clusters ([Figures S3B](#) and [S3D](#)). This is expected, as the cocktail of antibodies against presynaptic active-zone proteins should label the presynaptic terminals of both excitatory and inhibitory synapses, and excitatory presynaptic terminals would not be expected to pair with gephyrin. The population of small presynaptic clusters did not show any appreciable pairing with gephyrin clusters ([Figure S3B](#)).

These analyses demonstrated that we could identify synapses based on the size and signal density of the fluorescent clusters observed in STORM images. In the following experiments, we focus our analysis on the population of synaptic clusters with larger volumes. In contrast, similar analysis of either gephyrin or presynaptic clusters observed in the corresponding conventional images did not allow clear distinction between synaptic and non-synaptic clusters ([Figure S4A](#)).

Examination of the hundreds of thousands of automatically identified synapses in STORM images of the inner retina showed non-uniform distributions across the depth of the IPL ([Figure 2C](#)). The difference between gephyrin-paired (inhibitory) and unpaired (putative excitatory) presynaptic cluster intensities divided the IPL into several sublaminae, two of which coincided with On-Off DSGC stratification in sublaminae S3

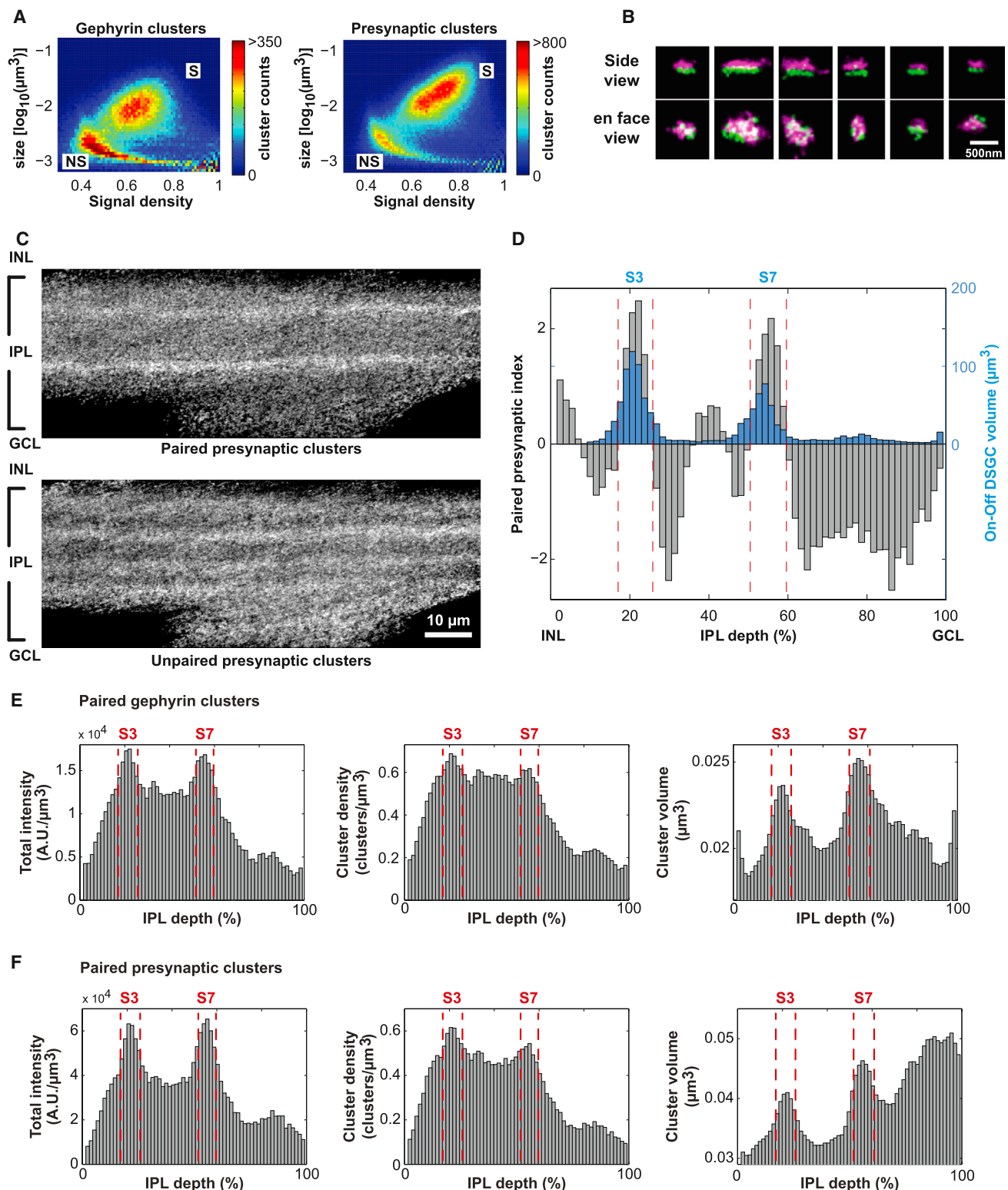


Figure 2. Automated Inhibitory Synapse Identification within the IPL

(A) Gephyrin (left) and presynaptic (right) clusters across the IPL can be separated into putative synaptic (S) and non-synaptic (NS) populations based on the volumes and signal densities of the clusters. Shown are the 2D distributions of cluster volume and signal density constructed from all gephyrin and presynaptic

(legend continued on next page)

and S7 (Figure 2D) (Vaney et al., 2012). Interestingly, the gephyrin and presynaptic signal intensities and the density and volumes of these gephyrin-positive inhibitory synapses all peaked in S3 and S7 (Figures 2E and 2F). As the size of inhibitory synapses correlates with synaptic strength (Lim et al., 1999; Nusser et al., 1997, 1998), this observation suggests that the inhibitory synapses subserving On-Off direction-selectivity may be among the strongest inhibitory connections in the mouse retina.

Identifying Inhibitory Synaptic Inputs to Labeled Neurons

To demonstrate the ability of our super-resolution platform to segment and analyze synaptic inputs onto identified neurons, we reconstructed two types of retinal ganglion cells and their associated inhibitory synaptic fields. Each of these datasets consisted of both STORM images (Figure 3A, left) and, for comparison, the corresponding conventional images (Figure 3A, right). To identify synaptic inputs onto neurons, we measured the density of gephyrin clusters and associated presynaptic signals as a function of distance to the neuron surface. Both density functions derived from STORM images were sharply peaked near the neuron surface with the gephyrin peak slightly inside the neuron and the presynaptic signal slightly outside the neuron as expected for input synapses (Figure 3B). These density peaks, in particular the gephyrin peak, were followed by a depletion zone, where the density dropped below the mean density of the surrounding IPL. For automated assignment of synapses to the neuron, we set a cutoff at the point where the gephyrin density dropped below the mean density of the surrounding IPL and selected only those gephyrin clusters located at a distance below the cutoff as synaptic inputs to the neuron (Figures 3B and 3D, left, and Movie S1).

Figure 3E and Movie S2 show the 1,017 inhibitory synapses assigned to a reconstructed On-Off DSGC. The number of synapses that we identified by STORM reconstruction here was similar to that estimated by previous EM reconstructions of SAC inputs to On-Off DSGCs (Briggman et al., 2011). Moreover, more than 98% of the synaptic gephyrin clusters assigned to the neuron had an apposing presynaptic partner. All of the gephyrin-presynaptic pairs assigned to the neuron were spatially oriented with the presynaptic structure more distant from the neuron than the postsynaptic structure (Figure 3C), which is consistent with these structures being input synapses onto the neuron. Together, these results further demonstrated the high accuracy in our synapse identification and assignment.

In comparison, assignment of synapses to neurons based on the corresponding conventional fluorescence images was less precise as the diffraction-limited resolution made it difficult both to identify synaptic clusters and also to set a proper cutoff value for assigning clusters to the neuron (Figure S4). As a result, this analysis resulted in substantial error rates (up to ~50%), depending on the selected cutoff distance (Figures 3D and S4C).

Distribution of Inhibitory Inputs to On-Off DSGCs

We next evaluated the size and position of all gephyrin-positive synapses within the dendritic arbor of each reconstructed cell. On-Off DSGCs, such as those shown in Figure 4A, exhibited non-random synapse distributions on both local and whole-cell scales (Figures 4B–4E). A Ripley's clustering analysis showed that synapses were significantly more depleted within ~1 μm of another synapse than would be predicted by a random distribution on the dendritic arbor (Figures 4B and S5A), likely reflecting a minimum inter-synapse spacing imposed by the finite size of each synapse, which is consistent with a previous observation (Bleckert et al., 2013). On the whole-cell scale, On-Off DSGCs exhibited sublaminae specificity with substantially higher synapse density in sublaminae S3 and S7 than in other sublaminae, even after normalization for the different surface areas of dendrites across the IPL depth (Figures 4C and S5B). This pattern is consistent with the specific innervation of On-Off DSGCs by SACs (Vaney et al., 2012), which also stratify in S3 and S7.

Distribution of Inhibitory Inputs to a Small-Field On-Center RGC

For comparison, we examined the sizes and spatial distribution of inhibitory synapses (936 total) onto a small-field On-center RGC (Figures 4F and 4G–4J and Movie S3), a putative type G6 as previously classified (Völgyi et al., 2009). Similar to On-Off DSGCs, synapses on this cell also exhibited a non-random spatial distribution on the local scale where Ripley's clustering analysis showed an ~1–2 μm depletion zone in the vicinity of each synapse (Figure 4G). However, the inhibitory synaptic input field of this neuron exhibited less sublaminae specificity than On-Off DSGCs on a whole-cell scale (Figure 4H).

Receptor Identity of Inhibitory Inputs to On-Off DSGCs

To demonstrate the capability of this super-resolution fluorescence reconstruction platform to determine the molecular identities of synaptic connections within neural circuits, we performed experiments to disambiguate different inhibitory synaptic input classes (GABAergic versus glycinergic) onto identified

clusters identified in the image block, with the cluster volume plotted on the log scale. The signal density is defined as the fraction of the volume occupied by the cluster that is positive for the gephyrin or presynaptic signal.

(B) Six example pairs of gephyrin (green) and presynaptic (magenta) clusters. The synapses are rotated to show side and en face views.

(C) Projection images of the IPL showing gephyrin-paired and unpaired synaptic clusters in the presynaptic channel. GCL: ganglion cell layer; INL, inner nuclear layer.

(D) The laminar distributions of the gephyrin-paired and unpaired presynaptic clusters divide the IPL into several sublaminae. Gray bars: presynaptic pairing index as a function of IPL depth. "Presynaptic pairing index" is calculated as the difference of the paired and unpaired presynaptic laminar intensity distributions after first standardizing each distribution to have a mean of zero and a SD of one. Blue bars: volume of the On-Off DSGC (μm^3) per 0.5 μm bin as a function of IPL depth. (E) The total signal intensity (left), average cluster density (middle), and average cluster volume (right) for each cubic micron of imaged tissue measured as a function of depth within the IPL for gephyrin clusters that are paired with presynaptic clusters.

(F) Similar to (E) but for presynaptic clusters that are paired with gephyrin clusters.

The delineation of sublaminae S3 and S7 in (E) and (F) was determined based on (D). See also Figure S3.

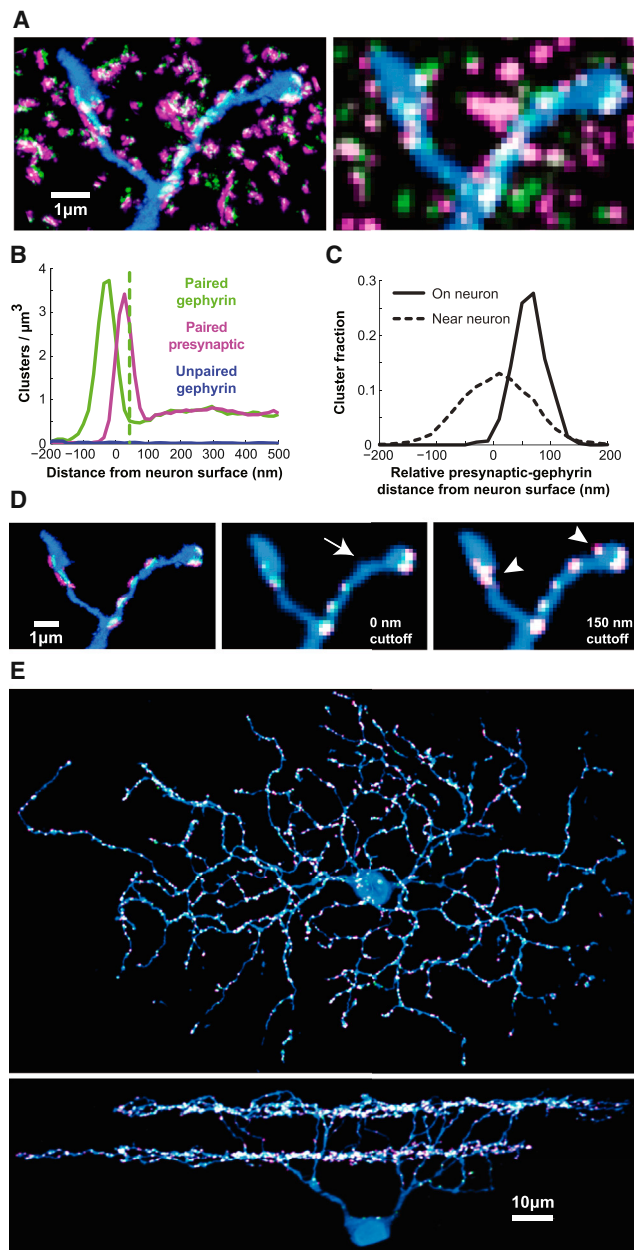


Figure 3. Automated Segmentation of Synaptic Inputs to Neurons

(A) STORM maximum projection image of a region containing a dendritic branch of a reconstructed On-Off DSGC (left) and the corresponding conventional image (right). Neurite is in blue, gephyrin in green, and presynaptic channel in magenta.

(B) The densities of the gephyrin clusters that are paired with presynaptic clusters (green trace), the unpaired gephyrin clusters (blue trace), and the gephyrin-paired presynaptic signal (magenta trace) measured as a function of the distance to the neuron surface. The distance at which the density peak of gephyrin clusters drops below the mean synapse density of the surrounding IPL (dashed green line) is used as a cutoff for defining gephyrin clusters on the neuron.

(C) For each synapse, we measured the distances of the presynaptic and postsynaptic signal to the neuron surface and defined the difference between these two distances as the relative presynaptic-gephyrin distance from the neuron surface. All synapses assigned to the On-Off DSGC show positive

relative distance values (solid line), which is consistent with these pairs being input synapses onto the neuron. In contrast, the spatial arrangement of nearby synapses within 500 nm of the neuron (dashed line) shows a broad distribution of both positive and negative relative distance values, indicating a random orientation of nearby synapses with respect to the neuron surface.

(D) Assignment of synapses in the STORM image based on the cutoff selected in (B) reveals adjacent presynaptic and postsynaptic structures associated with the neuron (left). In contrast, assignment of synapses in the conventional images with a cutoff at 0 nm (middle) or 150 nm (right) show false-negative (arrows) and false-positive synapse assignments (arrowheads).

(E) En face view (top) and side view (bottom) of the STORM maximum intensity projection of a reconstructed On-Off DSGC (blue) with associated synaptic gephyrin (green) and presynaptic (magenta) clusters. Although gephyrin and presynaptic clusters are clearly resolved in the original reconstruction (Figure 3D and Movie S1), they appear as overlapping white dots here due to image downsampling.

neurons. We labeled retinæ with either an antibody against the $\alpha 2$ subunit of the GABA(A) receptor (GABA(A)R $\alpha 2$) or an antibody cocktail against all alpha subunits of glycine receptors (GlyR $\alpha 1-4$), in addition to antibodies against GFP and gephyrin for marking neurons and inhibitory synapses, respectively. To determine whether each gephyrin-positive inhibitory synapse contained GABA(A)R $\alpha 2$ or glycine receptors, we examined whether the corresponding synaptic gephyrin cluster was paired with a specific receptor cluster by using the same approach described above for pairing presynaptic and postsynaptic structures (Figures S6A and S6B).

In GABA(A)R $\alpha 2$ -labeled samples, On-Off DSGC dendrites contained many GABA(A)R $\alpha 2$ -paired gephyrin clusters but strikingly rare unpaired gephyrin clusters (Figures 5A, 5B, and S6C). The gephyrin and GABA(A)R $\alpha 2$ signal intensities in these synapses were strongly correlated with a Pearson coefficient of 0.82 (Figure S6E), suggesting that gephyrin intensity in these synapses correlates with synaptic strength, as is the case elsewhere in the nervous system (Lim et al., 1999; Nusser et al., 1997, 1998). Quantitatively, $97\% \pm 1\%$ of the gephyrin-positive synapses on On-Off DSGCs contained GABA(A)R $\alpha 2$, suggesting a high labeling efficiency of the receptors. Compared with synapses on On-Off DSGCs, only $\sim 45\%$ of all gephyrin-positive synapses analyzed across the IPL contained GABA(A)R $\alpha 2$, demonstrating a strong enrichment of GABA(A)R $\alpha 2$ in the synapses onto On-Off DSGCs. Although not all GABA receptor types are anchored at synapses by a gephyrin scaffold (Brickley and Mody, 2012; Tretter et al., 2012; Tyagarajan and Fritschy, 2014), gephyrin-independent GABA receptors are unlikely to contribute to direction selectivity (Brickley and Mody, 2012; Massey et al., 1997). This, together with the similar synapse counts observed between our experiments and previous EM reconstructions of SAC synapses onto On-Off DSGCs (Briggman et al., 2011), suggests that the vast majority, if not all, of the inhibitory synapses onto On-Off DSGCs are gephyrin positive. Hence, our observations suggest that nearly all of the inhibitory synapses onto On-Off DSGCs contain the GABA(A)R $\alpha 2$ subunit.

In stark contrast to the GABA(A)R $\alpha 2$ -labeled samples, in the GlyR $\alpha 1-4$ labeled samples, we observed very few GlyR $\alpha 1-4$ positive synaptic gephyrin clusters on On-Off DSGCs (Figures 5C, 5D, and S6D). Quantitatively, only $8\% \pm 4\%$ of the synaptic

See also Figure S4 and Movies S1 and S2.

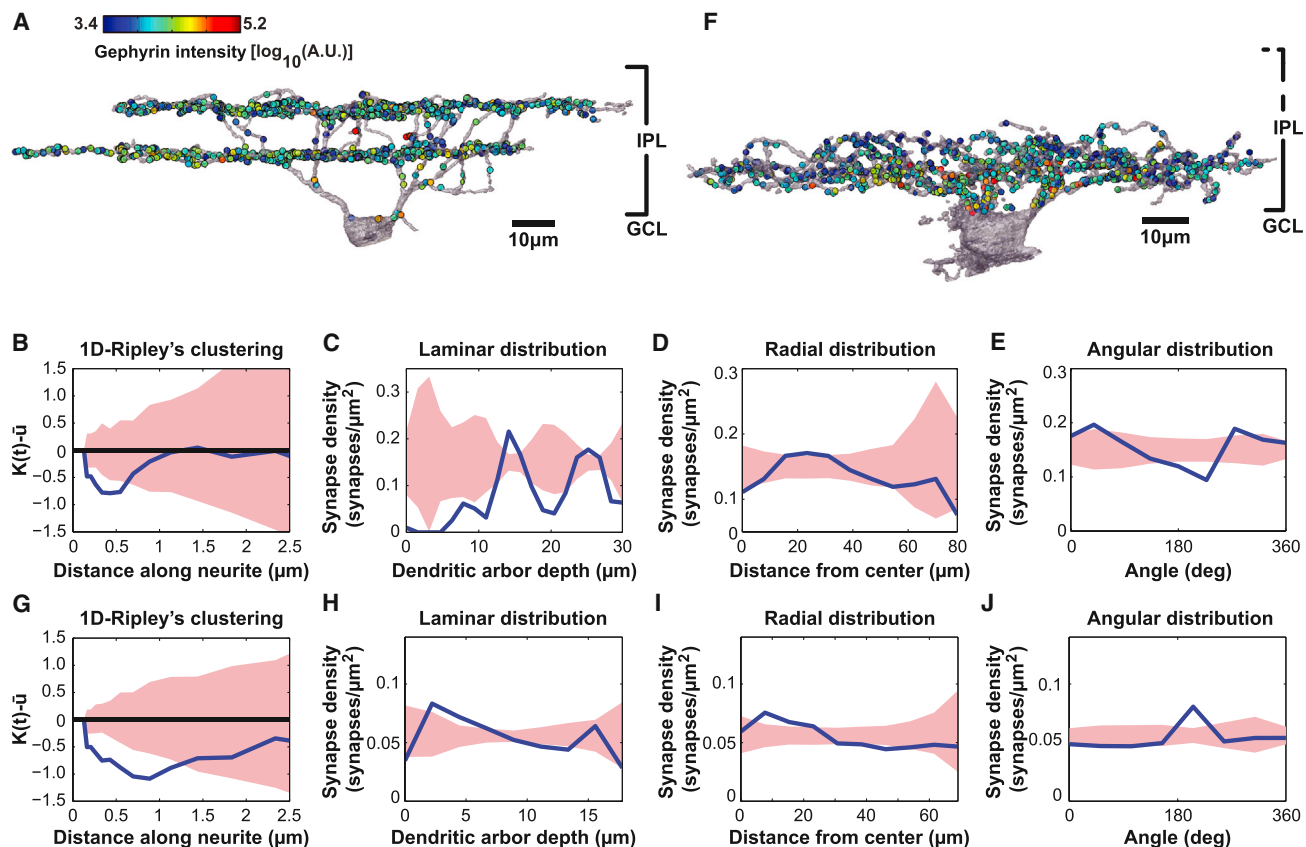


Figure 4. Distributions of Inhibitory Synapses on an On-Off DSGC and a Small-Field On-Center RGC

(A) Surface renderings of the On-Off DSGC (gray) shown in Figure 3E with all inhibitory synaptic inputs marked by circles whose color (blue to red) reflects gephyrin cluster intensity on a log scale.

(B) A one-dimensional Ripley's clustering analysis along the path of the skeletonized neuron. Negative value of the Ripley's function $K(t) - \bar{u}$ at short inter-synaptic distances indicate that, near any given synapse, the density of other synapses is significantly lower than a random distribution (see [Experimental Procedures](#) for the definition of Ripley's K function).

(C–E) The laminar (C), radial (D), and angular (E) distributions of the inhibitory synapse densities on the On-Off DSGC.

(F) Surface renderings of a small-field On-center RGC (gray) with all inhibitory synaptic inputs marked by circles whose color (blue to red) reflects gephyrin cluster intensity.

(G–J) Similar to (B–E) but for the On-center RGC.

Pink regions in (B–E) and (G–J) reflect 5/95% confidence intervals of random distributions derived from 1,000 randomizations of the synapse positions.

See also [Figure S5](#) and [Movies S2](#) and [S3](#).

gephyrin clusters on On-Off DSGCs contained any GlyR α 1–4 signal, and even these synapses exhibited extremely sparse GlyR α 1–4 labeling relative to nearby glycine-positive synapses not on the labeled On-Off DSGCs (Figures S6F and S6G). Since these nearby synapses contained substantial GlyR α 1–4 signal, the lack of GlyR α 1–4 in the On-Off DSGC synapses could not be attributed to low receptor labeling efficiency. Moreover, while previous work has shown a strong correlation between glycine receptor and gephyrin expression at synapses (Specht et al., 2013), we observed little correlation between the intensity of gephyrin and GlyR signals for these GlyR-positive gephyrin clusters on the On-Off DSGCs (Figure S6E). These results suggest that these sparse, low-intensity GlyR punctae probably reflect non-specific background labeling, and even if they were specific synaptic labeling, they would contribute relatively little synaptic current due to the low receptor abundance. As gephyrin is

required for clustering glycine receptors at synapses (Feng et al., 1998; Fischer et al., 2000; Kirsch et al., 1993), our results thus indicate that On-Off DSGCs in the mouse retina receive little monosynaptic glycinergic input.

In contrast to the STORM results, analysis of the corresponding conventional fluorescence images showed that a substantial population (20%–30%) of the gephyrin-labeled “synapses” assigned to On-Off DSGCs were GABA(A)R α 2 negative (Figure S7). These errors arise primarily from two sources: (1) it is difficult to separate synaptic gephyrin clusters from non-specific background labeling or trafficking vesicles containing gephyrin based on conventional images (Figure S4A), and hence some of the gephyrin clusters assigned to the neuron may not correspond to synapses; and (2) synapses near the neuron, but not on the neuron, can be mistakenly assigned to the neuron because of the limited resolution of the conventional images (Figure S4C).

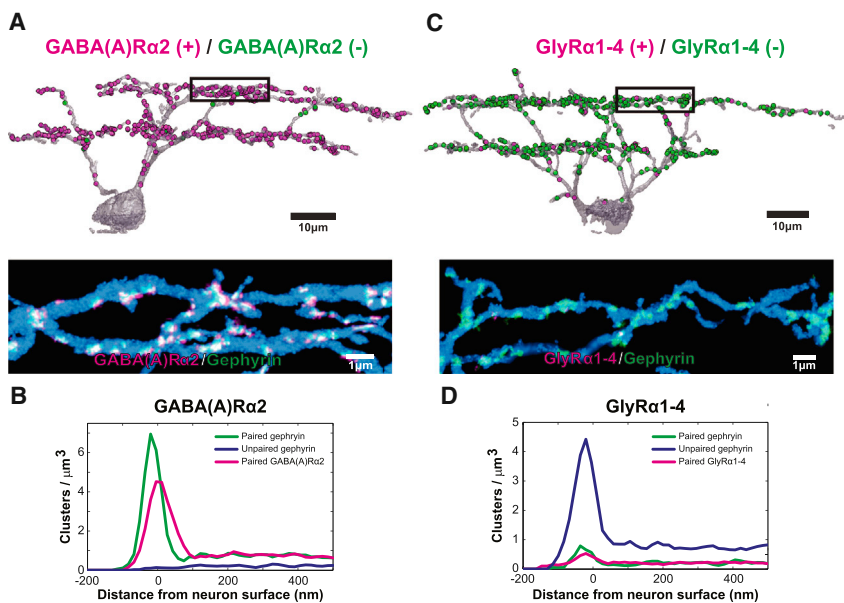


Figure 5. Receptor Identity of the Inhibitory Synaptic Inputs to On-Off DSGCs

(A) Top: surface rendering of a central cross-section of an On-Off DSGC (gray) with GABA(A)Rα2-positive (+) and GABA(A)Rα2-negative (-) inhibitory synapses marked as magenta and green circles, respectively. STORM image of the boxed region is shown in the bottom panel. Neuron: blue. Gephyrin: green. GABA(A)Rα2: magenta.

(B) The GABA(A)Rα2-paired gephyrin (green), gephyrin-paired GABA(A)Rα2 (magenta), and unpaired gephyrin cluster (blue) densities as a function of the distance to the neuron shown in (A).

(C) Top: surface rendering of a central cross-section of an On-Off DSGC (gray) with GlyRα1-4-positive (+) and GlyRα1-4-negative (-) inhibitory synapses marked as magenta and green circles, respectively. STORM image of the boxed region is shown in the bottom panel. Neuron: blue. Gephyrin: green. GlyRα1-4: magenta.

(D) The GlyRα1-4-paired gephyrin (green), gephyrin-paired GlyRα1-4 (magenta), and unpaired gephyrin cluster (blue) densities as a function of the distance to the neuron shown in (C).

See also Figures S6 and S7.

Inhibitory Inputs and Outputs of a Glycinergic Interneuron

Last, we imaged gephyrin-positive inhibitory synapses associated with a subtype of narrow-field amacrine cell (NFAC) (Figure 6A and Movie S4), putatively a Type 7 based on previous characterization (Pang et al., 2012). NFACs mediate crossover inhibition between On and Off sublaminae of the IPL via glycinergic inhibition (Werblin, 2010). In contrast to On-Off DSGCs, the surface of this NFAC was highly enriched with paired GlyRα1-4 and gephyrin clusters but largely depleted of unpaired, GlyRα1-4-negative gephyrin clusters (Figures 6A and 6B). The resolution of STORM allowed us to visualize the orientations of gephyrin-receptor pairs relative to the neuron surface and determine whether these structures were input synapses onto the cell or output synapses from the cell (Figure 6C). Unlike GABAergic synapses onto On-Off DSGCs, which were all input synapses (Figures 6D and 3C), the glycinergic synapses on the NFAC contain both input and output synapses (Figures 6C and 6D). Both synaptic inputs and outputs exhibited sublamina specificity with enrichment in the Off sublaminae (Figure 6E), suggestive of this cell being an On-center responsive Type 7 glycinergic amacrine cell (Pang et al., 2012) providing crossover inhibitory output to the Off sublaminae (Werblin, 2010).

About 85% of the synapses on this neuron contained glycine receptors (Figure 6E), again indicating a high receptor labeling efficiency in our samples. Since NFACs are glycinergic cells, it is not surprising that the observed output synapses from this cell were mostly GlyRα1-4 positive. It is, however, interesting to observe that the majority of gephyrin-positive inputs onto this cell were also GlyRα1-4 positive, suggesting that this type of NFAC receives inhibitory input signals mainly from other glycinergic amacrine cells, though our results do not exclude the possibility that this cell type also receives some GABAergic inputs.

DISCUSSION

Mapping the spatial organization and molecular identity of synaptic connections within neuronal networks is important for understanding how the nervous system functions. Here, we developed a super-resolution platform for volumetric reconstruction and automated segmentation of endogenous molecular targets in tissue and demonstrated the ability of this platform to identify the spatial patterns and molecular identity of inhibitory synapses within neuropil, as well as onto individual neurons using the mouse retina as a model system.

This method provides several benefits for reconstructing synaptic connectivity. First, the superior resolution of this approach, as compared to conventional fluorescence imaging, allows more accurate identification of synapses and assignment of synapses to neurons. Indeed, when comparing results from the same tissue samples, we found that conventional fluorescence imaging led to substantial errors both in the identification of synapses and in the assignment of synapses to neurons even with the improved z resolution afforded by ultrathin sectioning. These errors resulted in misidentification of inhibitory synaptic types onto On-Off DSGCs, which could lead to substantial misinterpretation of cellular physiology. In addition, the resolution provided by STORM also allowed us to quantitatively measure synapse size, which is often a good indicator of synaptic strength (Nusser et al., 1997, 1998). This ability allowed us to map the relative strengths of inhibitory synapses at different locations both on identified neurons and across the IPL.

A second benefit of the super-resolution reconstruction platform is its ability to use standard immunohistochemistry for labeling multiple endogenous protein targets of interest, which allows the determination of the molecular identities of synapses. Such information is difficult to ascertain using EM reconstructions alone but is important for interpreting the function of specific synapses in neural circuits (Bargmann and Marder, 2013).

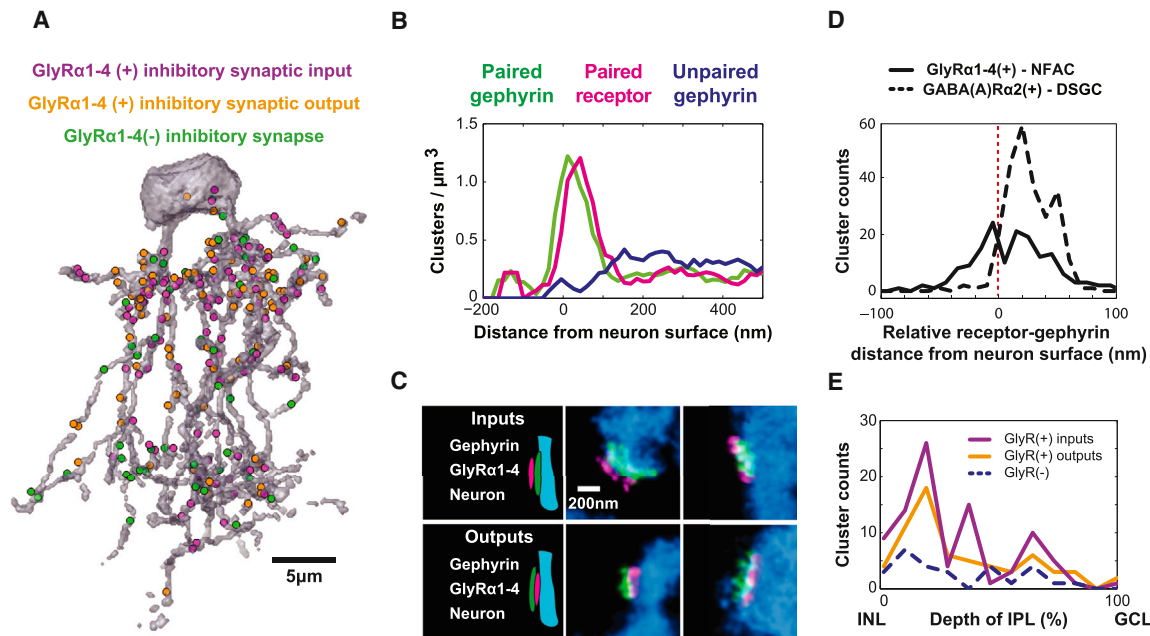


Figure 6. Input and Output Inhibitory Synapses of a NFAC

(A) Surface rendering of a NFAC (gray) with GlyR α 1-4 (+) input (purple circles), output (orange circles), and GlyR α 1-4 (-) synapses (green circles) shown. (B) The GlyR α 1-4-paired gephyrin (green), gephyrin-paired GlyR α 1-4 (magenta), and unpaired gephyrin cluster (blue) densities as a function of the distance to the surface of the neuron. (C) Examples of input and output synapses distinguished by the positions of receptor and gephyrin signals relative to the surface of the neuron. Input synapses have the gephyrin clusters (green) in the dendrites (blue) and receptor clusters (magenta) on the surface. Output synapses have the receptor cluster immediately adjacent to the dendrite surface and the gephyrin clusters farther outside. (D) The relative displacement of receptor and gephyrin clusters from the neuron surface, with positive values indicating receptor being farther from the neuron (input synapses) and negative value indicating gephyrin being farther from the neuron (output synapses). The solid line shows the distribution for glycinergic synapses associated with the NFAC, and the dashed line shows the distribution for the GABAergic synapses associated with an On-Off DSGC. (E) Laminar distributions of the input (purple) and output (orange) GlyR α 1-4-positive synapses and the GlyR α 1-4-negative synapses (blue) on the NFAC. See also [Movie S4](#).

Taking advantage of this capability, we showed that gephyrin-positive inhibitory synapses onto On-Off DSGCs were overwhelmingly GABAergic and each contained the GABA(A) receptor α 2 subunit, suggesting that this receptor subunit is important for generating postsynaptic currents during motion detection. This result is consistent with previous data showing the enrichment of GABA(A) α 2 in On-Off DSGC synapses and reduction in direction-selective responses in the GABA(A) α 2 knockout mouse (Auerkorte et al., 2012). Our reconstructions also showed that On-Off DSGCs receive little, if any, monosynaptic glycinergic input. These structural data, together with the observations that blocking GABA receptors largely eliminates inhibitory currents in On-Off DSGCs (Stafford et al., 2014; Trenholm et al., 2011), suggest that glycinergic modulation of On-Off DSGCs does not occur via direct glycinergic inputs onto these neurons but likely through glycinergic inhibition of bipolar cells or SACs that are presynaptic to On-Off DSGCs (Ishii and Kaneda, 2014; Majumdar et al., 2009; Zhang and McCall, 2012).

A third strength of this reconstruction platform is its ability to perform automated segmentation of synaptic connections in neural circuits without manual annotation. This automated analysis capability greatly speeds up the image processing required to extract biological information from individual reconstructions.

For example, the image processing for volumetric reconstruction and segmentation of a whole On-Off-DSGC cell and associated synapses took <3 days of computation time without any need for manual segmentation or correction. In this work, the rate-limiting step of our reconstructions was the STORM image acquisition time, as imaging an entire On-Off-DSGC of $2.3 \times 10^5 \mu\text{m}^3$ in four color channels took ~3 weeks using a STORM setup equipped with an EMCCD camera. Our recent switch to a scientific CMOS (sCMOS) camera with a larger field of view and higher frame rate (Huang et al., 2013) shortened the imaging time of a comparable volume to ~3 days. We envision this automated imaging and segmentation pipeline to be beneficial for determining neural circuit properties in different genetic mutant and disease models or at different time points during development, where a large number of reconstructions are needed.

One potential limitation of this super-resolution fluorescence platform, as compared with EM approaches, is the density of neuronal processes that can be reconstructed within a volume. In this work, we reconstructed the spatial distributions and molecular identities of synapses onto individual neurons in Thy1-GFP/YFP transgenic mice, in which only sparse subsets of neurons are labeled. We expect that our approach can be extended to the reconstruction of multiple, synaptically coupled

neurons using recently developed high-density, high-antigenicity, genetic labeling approaches (Cai et al., 2013; Loulier et al., 2014; Viswanathan et al., 2015) or by microinjection of probes to directly label multiple neurons. Although the image resolution here was limited in the z direction by the 70 nm section thickness, we anticipate a substantial improvement in z resolution by using 3D STORM (Huang et al., 2008). In particular, using high-precision z-localization approaches (Jia et al., 2014; Shtengel et al., 2009; Xu et al., 2012), the optical resolution can reach ~10 nm in all three dimensions. However, this resolution is still lower than that achievable by EM, and the labeling density may impose an additional limitation on resolution. Together, these may limit the density of neurites that can be reconstructed, and it remains to be determined whether this STORM platform can be used for dense reconstruction of all neurons in a volume.

With its unique capabilities complementary to existing reconstruction methods, we expect that this volumetric super-resolution reconstruction platform will enable a variety of synaptic connectivity analyses that will substantially enhance our understanding of the structural basis of nervous system function. The ability to reconstruct and identify endogenous molecular targets in large tissue volumes should also benefit the studies of many other biological systems.

EXPERIMENTAL PROCEDURES

Animals

Animal work was performed in accordance with protocols approved by the Institutional Animal Care and Use Committee at Harvard University. Adult transgenic mice (Tg(Thy1-EGFP)MJrs/J or YFP (Tg(Thy1-YFP)HJrs/J, The Jackson Laboratory) (Feng et al., 2000), both male and female animals 6–24 weeks of age, were used in our experiments.

Retinal Tissue Preparation

Whole eye-cups were immersion fixed in 4% paraformaldehyde for 10–60 min at room temperature. Both whole-mount and vibratome-sectioned retinæ were used for labeling. For whole-mount labeling, retinæ were laid flat on nitrocellulose membranes, and individual labeled neurons were excised in circular punches (diameter ~500 µm, thickness ~200 µm). For vibratome section labeling, retinæ were immersed in 37°C 2%–3% agarose, cooled on ice, and sectioned at 50–150 µm thickness in 1× DPBS.

Immunohistochemistry

Retinæ were blocked in 10% normal donkey serum in 1× DPBS with 0.3% Triton X-100 and 0.02%–0.05% sodium azide for 2–3 hr at room temperature and incubated in primary antibody solutions diluted in blocking buffer overnight for 3–4 nights at 4°C. A complete list of all primary antibodies tested in this work is provided in Table S1 with the antibodies selected for the STORM reconstructions highlighted. Following primary antibody incubation, retinæ were washed 6 times for 20 min each in 2% normal donkey serum in 1× DPBS at room temperature and incubated in secondary antibodies (detailed in the Supplemental Experimental Procedures) overnight at 4°C for 1–2 nights to label the neuron with photoswitchable dye Atto 488 and two synaptic targets (gephyrin and presynaptic proteins or gephyrin and receptors) with photoswitchable dyes Alexa Fluor 647 and DyLight 750, respectively. The antibodies for labeling synaptic proteins were also conjugated to Alexa Fluor 405 to facilitate photoactivation of Alexa Fluor 647 and DyLight 750. Retinæ were then washed 6 times for 20 min each in 1× DPBS at room temperature and incubated overnight in Cy3B-labeled WGA.

Postfixation, Dehydration, and Embedding in Epoxy Resin

Labeled retinæ were postfixed for 2 hr in 3% paraformaldehyde and 0.1% glutaraldehyde diluted in 1× DPBS. Postfixed retinæ were dehydrated in a

graded series of ethanol washes (50%/70%/90%/100% two times) for 10–20 min each and then incubated in UltraBed Epoxy Resin (Electron Microscopy Sciences) solutions of increasing concentration for 2 hr each (75% ethanol/25% resin; 50% ethanol/50% resin; 25% ethanol/75% resin; 100% resin 2 times). Dehydrated resin blocks were then polymerized in UltraBed overnight for 16 hr at 70°C.

Ultrathin Sectioning

Ultrathin sections were cut at 70 nm on a Leica UC7 Ultramicrotome (Leica Microsystems) using an ultra Jumbo diamond knife (Diatome). The section thickness was verified in two independent ways, as described in the Supplemental Experimental Procedures. Sections were collected on glass coverslips coated with 0.5% gelatin/0.05% chromium potassium sulfate. Coverslips were dried at 60°C for 25 min.

Preparation of Coverslips for Imaging

Coverslips of tissue sections were immersed in 10% sodium ethoxide solution for 5–20 min to etch the embedding resin for optimal photoswitching of dyes. Fluorescent beads (mixture of 540/560 and 715/755 FluoSpheres from Life Technologies, detailed in the Supplemental Experimental Procedures) were spotted on the coverslips as fiducial markers. Coverslips were secured to glass slide flow channels, filled with STORM imaging buffer (10% glucose/17.5 µM glucose oxidase/708 nM catalase/10 mM MEA/10 mM NaCl/200 mM Tris), and sealed with epoxy.

Imaging Setup

Imaging was performed through Olympus UPlanSApo 100× 1.4 NA oil-immersion objectives mounted on Olympus IX71 inverted microscopes with back optics arranged for oblique incident angle illumination. The microscope contained a custom pentaband dichroic and pentanotch filter (Chroma Technology Corp) and laser lines at 488/561/647/750 nm (detailed in the Supplemental Experimental Procedures) for excitation of Atto 488, Cy3B, Alexa Fluor 647, and DyLight 750, respectively. A 405 nm laser was used for reactivation of dyes. Images were acquired on an Andor iXon3 897 or 897Ultra EMCCD camera through a QV2 quadview image splitter (Photometrics). Each camera pixel corresponded to ~158 nm in sample space, and the total imaging field size was ~40 µm × 40 µm. Axial focus during imaging was maintained in an automated manner as described previously (Dempsey et al., 2011).

Automated Image Acquisition

Tissue sections and fiducial bead fields were initially located using a 4× objective. Regions of interest (ROIs) were subsequently identified with a 100× objective. The stage position coordinates for each ROI were determined, and the position list for all ROIs on a coverslip was then used to generate a master file that controlled laser illumination, camera activation, stage movement, AOTF control, and shutter sequences for automated STORM and conventional imaging. Each imaging session began with imaging of low-density bead fields by first exciting the 540/560 beads at 488 nm and detecting in the Alexa Fluor 647, Cy3B, and Atto 488 channels and then exciting the 715/755 beads at 752 nm and detecting in the DyLight 750 and Alexa Fluor 647 channels. These low-density bead images were used for chromatic aberration correction across different color channels.

Next, each ROI was imaged at the conventional resolution in each of the four color channels (DyLight 750, Alexa Fluor 647, Cy3B, and Atto 488). Next, images of the high-density bead field were acquired in each of the four color channels for (1) flat-field correction to compensate for non-uniform illumination across the field of view and (2) lens distortion correction at image field edges.

STORM imaging of individual ROIs was next performed in four color channels. For each ROI, the DyLight 750 channel was imaged for ~4K–4.5K frames at 30 Hz, the Alexa Fluor 647 channel was imaged for 6K–7K frames at 60 Hz, and the Cy3B and Atto 488 channels were each imaged for ~10K frames at 60 Hz. To ensure that overlapping regions in each montage were not bleached, STORM movies were collected in two passes for each ROI, each consisting of half the total number of frames described above.

STORM Image Analysis

STORM movies were analyzed to determine the positions of individual molecules using a DAOSTORM algorithm (Babcock et al., 2012; Holden et al., 2011). Molecule lists were rendered as 2D images with 15.8 nm pixel size, which is close to both our ~20 nm STORM image resolution and 1/10 of the camera pixel size. For consistency of analysis, the conventional images were up-sampled to 15.8 nm/pixel. Chromatic aberrations were corrected using the transformation maps generated from the low-density bead field images, and lens-induced optical distortions were corrected using transformation maps generated from the high-density bead field images, as detailed in the Supplemental Experimental Procedures.

Alignment of Multiple Image Tiles within Individual Sections

Each STORM image was aligned to the corresponding conventional image using two-dimensional cross-correlation (Guizar-Sicairos et al., 2008). For mosaic imaging, Scale-Invariant Feature Transformation (SIFT) (Lowe, 2004) was used to find points of similarity between overlapping regions in adjacent image tiles in the WGA channel and generate a rigid alignment transformation that was applied to the conventional and STORM images to stitch overlapping image tiles. On average, the residual offset in alignment between SIFT points of similarity in two adjacent image tiles was <40 nm.

Alignment of Serial Sections

Corresponding SIFT features between adjacent sections were used to determine a rigid linear transformation between sections, which was applied to all sections in the dataset to achieve a coarse, 3D rigid alignment of the data. Then, we applied elastic registration (Saalfeld et al., 2012) to further improve the alignment accuracy between adjacent sections while minimizing the global deformation of the entire image block. The warping transforms generated in these steps were applied to all conventional fluorescence and STORM channels.

Segmentation of STORM and Conventional Fluorescence Images

STORM images were first filtered using a mask generated from the conventional images to remove background and signals from occasional debris on the coverslip. To generate this mask, the signals in the conventional images were thresholded using the lower threshold of a two-level Otsu threshold method (Otsu, 1979) that divided the signals in our images into three classes with the lowest-intensity class representing the background, the highest intensity class representing neuronal and synaptic features, and the middle class representing other low-intensity signals above background. To identify the surface of the neuron, we smoothed the neuron signal with a Gaussian kernel with $\sigma = 47$ nm and then binarized the neuron signal using the lower threshold of the two-level Otsu threshold method. To identify fluorescent clusters in the gephyrin, presynaptic or receptor channels in the STORM images, we applied a 79 nm Gaussian convolution to the signal in the XY plane and an isometric Gaussian convolution (~1 voxel) in Z and used the lower threshold of the two-level Otsu threshold method to binarize the image and identify connected components in three dimensions. Additional separation of over-connected clusters was performed using a watershed transformation. Processing of conventional images was performed similarly, except that we binarized the conventional images based on the higher threshold of a two-level Otsu threshold.

Two-Dimensional Analysis to Separate Different Populations of Gephyrin and Presynaptic Clusters

To determine whether a given cluster was synaptic, two parameters were considered for each cluster in the gephyrin and presynaptic channels: the volume of the cluster was calculated from the connected components within the segmented image. Second, the signal density was measured as the fraction of volume of the connected components that was occupied by signal-positive voxels in the raw data. For STORM images, plotting the distribution of these two parameters constructed from all clusters in the dataset as a 2D histogram showed two peaks. Separation of the two populations is described in the Supplemental Experimental Procedures.

Ripley's K Function

The Ripley's K function is calculated as $K(t) = \lambda^{-1} \sum_{i \neq j} I(d_{ij} < t) / n$, where t is the distance along neurites, λ is the average density of synapses on the neuron skeleton, I is the indicator function, d_{ij} is the distance between the i^{th} and j^{th} synapses, and n is the number of synapses on the neuron. \bar{t} is the average of $K(t)$ derived from 1,000 randomizations of synapse positions on the surface of the dendritic arbor.

A detailed complete description of the experimental procedures can be found in the Supplemental Experimental Procedures accompanying this paper.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, one table, and four movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.08.033>.

AUTHOR CONTRIBUTIONS

Y.M.S., C.M.S., and X.Z. designed the experiments. Y.M.S. and C.M.S. developed the super-resolution reconstruction platform and performed the imaging experiments. H.P.B. wrote software for automated image acquisition and analysis for constructing the STORM images. C.M.S. developed antibody-labeling strategies and performed STORM sample preparation. Y.M.S. developed image alignment, segmentation, and analysis algorithms and performed analysis of STORM datasets. C.M.S., Y.M.S., and X.Z. wrote the paper with input from H.P.B.

ACKNOWLEDGMENTS

We thank Daisy Spear, Mariah Evarts, and Sara Haddad for assistance with animals and Stephen Turney for providing custom confocal imaging equipment. We thank Joshua Sanes, Arjun Krishnaswamy, Melanie Samuel, Jeremy Kay, and the members of the X.Z. lab for discussions during the course of this project. This work was supported in part by Collaborative Innovation Awards of Howard Hughes Medical Institute (HHMI), the National Institute of Mental Health Conte Center Program (P50MH094271), and the Army Research Office MURI program (6019269). X.Z. is a HHMI investigator.

Received: April 18, 2015

Revised: July 22, 2015

Accepted: August 12, 2015

Published: October 1, 2015

REFERENCES

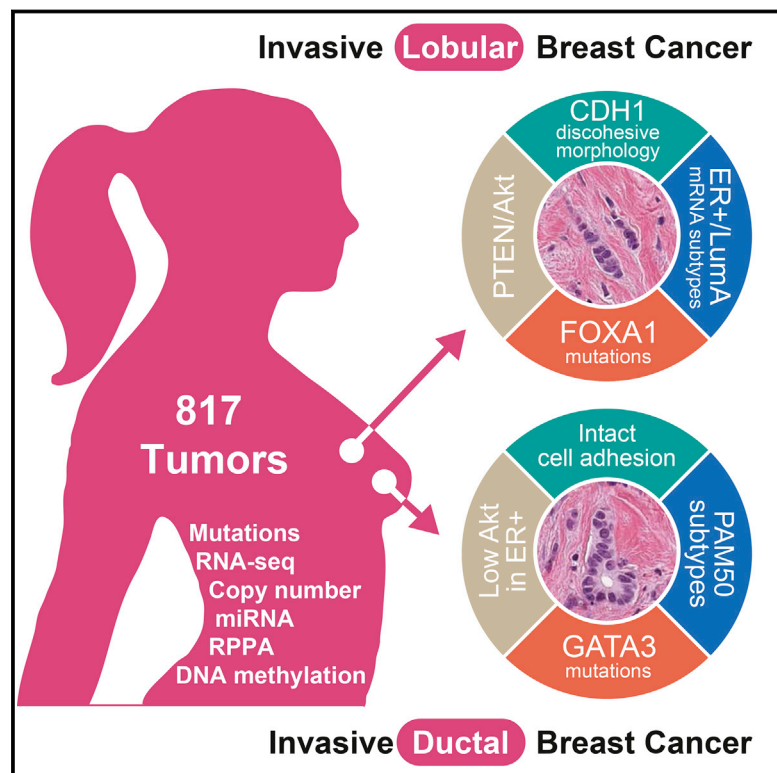
- Alivisatos, A.P., Chun, M., Church, G.M., Deisseroth, K., Donoghue, J.P., Greenspan, R.J., McEuen, P.L., Roukes, M.L., Sejnowski, T.J., Weiss, P.S., and Yuste, R. (2013). Neuroscience. The brain activity map. *Science* 339, 1284–1285.
- Anderson, J.R., Jones, B.W., Watt, C.B., Shaw, M.V., Yang, J.H., Demill, D., Lauritzen, J.S., Lin, Y., Rapp, K.D., Mastronarde, D., et al. (2011). Exploring the retinal connectome. *Mol. Vis.* 17, 355–379.
- Auferkorte, O.N., Baden, T., Kaushalya, S.K., Zabouri, N., Rudolph, U., Haverkamp, S., and Euler, T. (2012). GABA(A) receptors containing the $\alpha 2$ subunit are critical for direction-selective inhibition in the retina. *PLoS ONE* 7, e35109.
- Babcock, H., Sigal, Y.M., and Zhuang, X. (2012). A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Opt. Nanoscopy* 1, 1.
- Bargmann, C.I., and Marder, E. (2013). From the connectome to brain function. *Nat. Methods* 10, 483–490.
- Blecker, A., Parker, E.D., Kang, Y., Pancaroglu, R., Soto, F., Lewis, R., Craig, A.M., and Wong, R.O. (2013). Spatial relationships between GABAergic and glutamatergic synapses on the dendrites of distinct types of mouse retinal ganglion cells across development. *PLoS ONE* 8, e69612.

- Brickley, S.G., and Mody, I. (2012). Extrasynaptic GABA(A) receptors: their function in the CNS and implications for disease. *Neuron* 73, 23–34.
- Briggman, K.L., Helmstaedter, M., and Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. *Nature* 471, 183–188.
- Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W., and Sanes, J.R. (2013). Improved tools for the Brainbow toolbox. *Nat. Methods* 10, 540–547.
- Caldwell, J.H., Daw, N.W., and Wyatt, H.J. (1978). Effects of picrotoxin and strychnine on rabbit retinal ganglion cells: lateral interactions for cells with more complex receptive fields. *J. Physiol.* 276, 277–298.
- Chklovskii, D.B., Vitaladevuni, S., and Scheffer, L.K. (2010). Semi-automated reconstruction of neural circuits using electron microscopy. *Curr. Opin. Neurobiol.* 20, 667–675.
- Dani, A., Huang, B., Bergan, J., Dulac, C., and Zhuang, X. (2010). Superresolution imaging of chemical synapses in the brain. *Neuron* 68, 843–856.
- Dempsey, G.T., Vaughan, J.C., Chen, K.H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* 8, 1027–1036.
- Feng, G., Tintrup, H., Kirsch, J., Nichol, M.C., Kuhse, J., Betz, H., and Sanes, J.R. (1998). Dual requirement for gephyrin in glycine receptor clustering and molybdoenzyme activity. *Science* 282, 1321–1324.
- Feng, G., Mellor, R.H., Bernstein, M., Keller-Peck, C., Nguyen, Q.T., Wallace, M., Nerbonne, J.M., Lichtman, J.W., and Sanes, J.R. (2000). Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* 28, 41–51.
- Fischer, F., Kneussel, M., Tintrup, H., Haverkamp, S., Rauen, T., Betz, H., and Wässle, H. (2000). Reduced synaptic clustering of GABA and glycine receptors in the retina of the gephyrin null mutant mouse. *J. Comp. Neurol.* 427, 634–648.
- Fried, S.I., Münch, T.A., and Werblin, F.S. (2002). Mechanisms and circuitry underlying directional selectivity in the retina. *Nature* 420, 411–414.
- Gollisch, T., and Meister, M. (2010). Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron* 65, 150–164.
- Guizar-Sicairos, M., Thurman, S.T., and Fienup, J.R. (2008). Efficient subpixel image registration algorithms. *Opt. Lett.* 33, 156–158.
- Hell, S.W. (2007). Far-field optical nanoscopy. *Science* 316, 1153–1158.
- Helmstaedter, M. (2013). Cellular-resolution connectomics: challenges of dense neural circuit reconstruction. *Nat. Methods* 10, 501–507.
- Helmstaedter, M., Briggman, K.L., Turaga, S.C., Jain, V., Seung, H.S., and Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* 500, 168–174.
- Holden, S.J., Uphoff, S., and Kapanidis, A.N. (2011). DAOSTORM: an algorithm for high-density super-resolution microscopy. *Nat. Methods* 8, 279–280.
- Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319, 810–813.
- Huang, B., Babcock, H., and Zhuang, X. (2010). Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* 143, 1047–1058.
- Huang, F., Hartwich, T.M., Rivera-Molina, F.E., Lin, Y., Duim, W.C., Long, J.J., Uchil, P.D., Myers, J.R., Baird, M.A., Mothes, W., et al. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods* 10, 653–658.
- Ishii, T., and Kaneda, M. (2014). ON-pathway-dominant glycinergic regulation of cholinergic amacrine cells in the mouse retina. *J. Physiol.* 592, 4235–4245.
- Jain, V., Seung, H.S., and Turaga, S.C. (2010). Machines that learn to segment images: a crucial technology for connectomics. *Curr. Opin. Neurobiol.* 20, 653–666.
- Jensen, R.J. (1999). Responses of directionally selective retinal ganglion cells to activation of AMPA glutamate receptors. *Vis. Neurosci.* 16, 205–219.
- Jia, S., Vaughan, J.C., and Zhuang, X. (2014). Isotropic 3D Super-resolution Imaging with a Self-bending Point Spread Function. *Nat. Photonics* 8, 302–306.
- Kirsch, J., Wolters, I., Triller, A., and Betz, H. (1993). Gephyrin antisense oligonucleotides prevent glycine receptor clustering in spinal neurons. *Nature* 366, 745–748.
- Kittila, C.A., and Massey, S.C. (1995). Effect of ON pathway blockade on directional selectivity in the rabbit retina. *J. Neurophysiol.* 73, 703–712.
- Kleinfeld, D., Bharioke, A., Blinder, P., Bock, D.D., Briggman, K.L., Chklovskii, D.B., Denk, W., Helmstaedter, M., Kaufhold, J.P., Lee, W.C., et al. (2011). Large-scale automated histology in the pursuit of connectomes. *J. Neurosci.* 31, 16125–16138.
- Lichtman, J.W., and Denk, W. (2011). The big and the small: challenges of imaging the brain's circuits. *Science* 334, 618–623.
- Lim, R., Alvarez, F.J., and Walmsley, B. (1999). Quantal size is correlated with receptor cluster area at glycinergic synapses in the rat brainstem. *J. Physiol.* 516, 505–512.
- Loulier, K., Barry, R., Mahou, P., Le Franc, Y., Supatto, W., Matho, K.S., Ieng, S., Fouquet, S., Dupin, E., Benosman, R., et al. (2014). Multiplex cell and lineage tracking with combinatorial labels. *Neuron* 81, 505–520.
- Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 91–110.
- Majumdar, S., Weiss, J., and Wässle, H. (2009). Glycinergic input of widefield, displaced amacrine cells of the mouse retina. *J. Physiol.* 587, 3831–3849.
- Massey, S.C., Linn, D.M., Kittila, C.A., and Mirza, W. (1997). Contributions of GABAA receptors and GABAC receptors to acetylcholine release and directional selectivity in the rabbit retina. *Vis. Neurosci.* 14, 939–948.
- Micheva, K.D., and Smith, S.J. (2007). Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron* 55, 25–36.
- Micheva, K.D., Busse, B., Weiler, N.C., O'Rourke, N., and Smith, S.J. (2010). Single-synapse analysis of a diverse synapse population: proteomic imaging methods and markers. *Neuron* 68, 639–653.
- Miyawaki, A. (2015). Brain clearing for connectomics. *Microscopy (Oxf.)* 64, 5–8.
- Morgan, J.L., and Lichtman, J.W. (2013). Why not connectomics? *Nat. Methods* 10, 494–500.
- Nangneri, S., Flottmann, B., Horstmann, H., Heilemann, M., and Kner, T. (2012). Three-dimensional, tomographic super-resolution fluorescence imaging of serially sectioned thick samples. *PLoS ONE* 7, e38098.
- Nusser, Z., Cull-Candy, S., and Farrant, M. (1997). Differences in synaptic GABA(A) receptor number underlie variation in GABA mini amplitude. *Neuron* 19, 697–709.
- Nusser, Z., Hájos, N., Somogyi, P., and Mody, I. (1998). Increased number of synaptic GABA(A) receptors underlies potentiation at hippocampal inhibitory synapses. *Nature* 395, 172–177.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9, 62–66.
- Pang, J.J., Gao, F., and Wu, S.M. (2012). Physiological characterization and functional heterogeneity of narrow-field mammalian amacrine cells. *J. Physiol.* 590, 223–234.
- Punge, A., Rizzoli, S.O., Jahn, R., Wildanger, J.D., Meyer, L., Schönle, A., Kastrop, L., and Hell, S.W. (2008). 3D reconstruction of high-resolution STED microscope images. *Microsc. Res. Tech.* 71, 644–650.
- Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* 3, 793–795.
- Saalfeld, S., Fetter, R., Cardona, A., and Tomancak, P. (2012). Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nat. Methods* 9, 717–720.
- Shtengel, G., Galbraith, J.A., Galbraith, C.G., Lippincott-Schwartz, J., Gillette, J.M., Manley, S., Sougrat, R., Waterman, C.M., Kanchanawong, P., Davidson, M.W., et al. (2009). Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proc. Natl. Acad. Sci. USA* 106, 3125–3130.

- Specht, C.G., Izeddin, I., Rodriguez, P.C., El Beheiry, M., Rostaing, P., Darzacq, X., Dahan, M., and Triller, A. (2013). Quantitative nanoscopy of inhibitory synapses: counting gephyrin molecules and receptor binding sites. *Neuron* 79, 308–321.
- Stafford, B.K., Park, S.J., Wong, K.Y., and Demb, J.B. (2014). Developmental changes in NMDA receptor subunit composition at ON and OFF bipolar cell synapses onto direction-selective retinal ganglion cells. *J. Neurosci.* 34, 1942–1948.
- Stasheff, S.F., and Masland, R.H. (2002). Functional inhibition in direction-selective retinal ganglion cells: spatiotemporal extent and intralaminar interactions. *J. Neurophysiol.* 88, 1026–1039.
- Trenholm, S., Johnson, K., Li, X., Smith, R.G., and Awatramani, G.B. (2011). Parallel mechanisms encode direction in the retina. *Neuron* 71, 683–694.
- Tretter, V., Mukherjee, J., Maric, H.M., Schindelin, H., Sieghart, W., and Moss, S.J. (2012). Gephyrin, the enigmatic organizer at GABAergic synapses. *Front. Cell. Neurosci.* 6, 23.
- Tyagarajan, S.K., and Fritschy, J.M. (2014). Gephyrin: a master regulator of neuronal function? *Nat. Rev. Neurosci.* 15, 141–156.
- Vaney, D.I., Sivyer, B., and Taylor, W.R. (2012). Direction selectivity in the retina: symmetry and asymmetry in structure and function. *Nat. Rev. Neurosci.* 13, 194–208.
- Viswanathan, S., Williams, M.E., Bloss, E.B., Stasevich, T.J., Speer, C.M., Nern, A., Pfeiffer, B.D., Hooks, B.M., Li, W.P., English, B.P., et al. (2015). High-performance probes for light and electron microscopy. *Nat. Methods* 12, 568–576.
- Völgyi, B., Chheda, S., and Bloomfield, S.A. (2009). Tracer coupling patterns of the ganglion cell subtypes in the mouse retina. *J. Comp. Neurol.* 512, 664–687.
- Wei, W., Hamby, A.M., Zhou, K., and Feller, M.B. (2011). Development of asymmetric inhibition underlying direction selectivity in the retina. *Nature* 469, 402–406.
- Werblin, F.S. (2010). Six different roles for crossover inhibition in the retina: correcting the nonlinearities of synaptic transmission. *Vis. Neurosci.* 27, 1–8.
- Xu, K., Babcock, H.P., and Zhuang, X. (2012). Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nat. Methods* 9, 185–188.
- Zhang, C., and McCall, M.A. (2012). Receptor targets of amacrine cells. *Vis. Neurosci.* 29, 11–29.

Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer

Graphical Abstract



Authors

Giovanni Ciriello, Michael L. Gatz, Andrew H. Beck, ..., Tari A. King, TCGA Research Network, Charles M. Perou

Correspondence

cperou@med.unc.edu

In Brief

A comprehensive analysis of 817 breast tumor samples determines invasive lobular carcinoma as a molecularly distinct disease with characteristic genetic features, providing key information for patient stratification that may allow a more informed clinical follow-up.

Highlights

- Invasive lobular carcinoma (ILC) is a clinically and molecularly distinct disease
- ILCs show CDH1 and PTEN loss, AKT activation, and mutations in *TBX3* and *FOXA1*
- Proliferation and immune-related gene expression signatures define 3 ILC subtypes
- Genetic features classify mixed tumors into lobular-like and ductal-like subgroups



Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer

Giovanni Ciriello,^{1,2,23} Michael L. Gatz,^{3,4,23} Andrew H. Beck,⁵ Matthew D. Wilkerson,⁶ Suhm K. Rhie,⁷ Alessandro Pastore,² Hailei Zhang,⁸ Michael McLellan,⁹ Christina Yau,¹⁰ Cyriac Kandoth,¹¹ Reanne Bowlby,¹² Hui Shen,¹³ Sikander Hayat,² Robert Fieldhouse,² Susan C. Lester,⁵ Gary M.K. Tse,¹⁴ Rachel E. Factor,¹⁵ Laura C. Collins,⁵ Kimberly H. Allison,¹⁶ Yunn-Yi Chen,¹⁸ Kristin Jensen,^{16,17} Nicole B. Johnson,⁵ Steffi Oesterreich,¹⁹ Gordon B. Mills,²⁰ Andrew D. Cherniack,⁸ Gordon Robertson,¹² Christopher Benz,¹⁰ Chris Sander,² Peter W. Laird,¹³ Katherine A. Hoadley,³ Tari A. King,²¹ TCGA Research Network,²² and Charles M. Perou^{3,*}

¹Department of Medical Genetics, University of Lausanne (UNIL), 1011 Lausanne, Switzerland

²Computational Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

³Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

⁴Rutgers Cancer Institute of New Jersey, New Brunswick, NJ 08903, USA

⁵Department of Pathology, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA, 02215, USA

⁶Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

⁷Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, 90033, USA

⁸The Eli and Edythe L. Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA

⁹The Genome Institute, Washington University School of Medicine, MO, 63108, USA

¹⁰Buck Institute For Research on Aging, Novato, CA, 94945, USA

¹¹Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

¹²Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, V5Z4S6, Canada

¹³Center for Epigenetics, Van Andel Research Institute, Grand Rapids, MI, 49503, USA

¹⁴Department of Anatomical and Cellular Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong

¹⁵Department of Pathology, School of Medicine, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA

¹⁶Department of Pathology, School of Medicine, Stanford University Medical Center, Stanford University, Stanford, CA, USA

¹⁷VA Palo Alto Healthcare System, Palo Alto, 94304, CA, USA

¹⁸Department of Pathology and Laboratory Medicine, University of California, San Francisco, CA, 94143, USA

¹⁹Department of Pharmacology and Chemical Biology, Women's Cancer Research Center, University of Pittsburgh Cancer Institute, Pittsburgh, PA, 15232, USA

²⁰MD Anderson Cancer Center, The University of Texas, Houston, TX, 77230, USA

²¹Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

²²<http://cancergenome.nih.gov/>

²³Co-first author

*Correspondence: cperou@med.unc.edu

<http://dx.doi.org/10.1016/j.cell.2015.09.033>

SUMMARY

Invasive lobular carcinoma (ILC) is the second most prevalent histologic subtype of invasive breast cancer. Here, we comprehensively profiled 817 breast tumors, including 127 ILC, 490 ductal (IDC), and 88 mixed IDC/ILC. Besides E-cadherin loss, the best known ILC genetic hallmark, we identified mutations targeting PTEN, TBX3, and FOXA1 as ILC enriched features. PTEN loss associated with increased AKT phosphorylation, which was highest in ILC among all breast cancer subtypes. Spatially clustered FOXA1 mutations correlated with increased FOXA1 expression and activity. Conversely, GATA3 mutations and high expression characterized luminal A IDC, suggesting differential modulation of ER activity in ILC and IDC. Proliferation and immune-related signatures determined three ILC transcriptional subtypes associated with survival differences. Mixed IDC/ILC cases were molecularly classified as ILC-like

and IDC-like revealing no true hybrid features. This multidimensional molecular atlas sheds new light on the genetic bases of ILC and provides potential clinical options.

INTRODUCTION

Invasive lobular carcinoma (ILC) is the second most frequently diagnosed histologic subtype of invasive breast cancer, constituting ~10%–15% of all cases. The classical form (Foote and Stewart, 1946) is characterized by small discohesive neoplastic cells invading the stroma in a single-file pattern. The discohesive phenotype is due to dysregulation of cell-cell adhesion, primarily driven by lack of E-cadherin (CDH1) protein expression observed in ~90% of ILCs (McCart Reed et al., 2015; Morrogh et al., 2012). This feature is the ILC hallmark, and immunohistochemistry (IHC) scoring for CDH1 expression is often used to discriminate between lesions with borderline ductal versus lobular histological features. ILC variants have also been described, yet all display loss of E-cadherin expression (Dabbs et al., 2013).

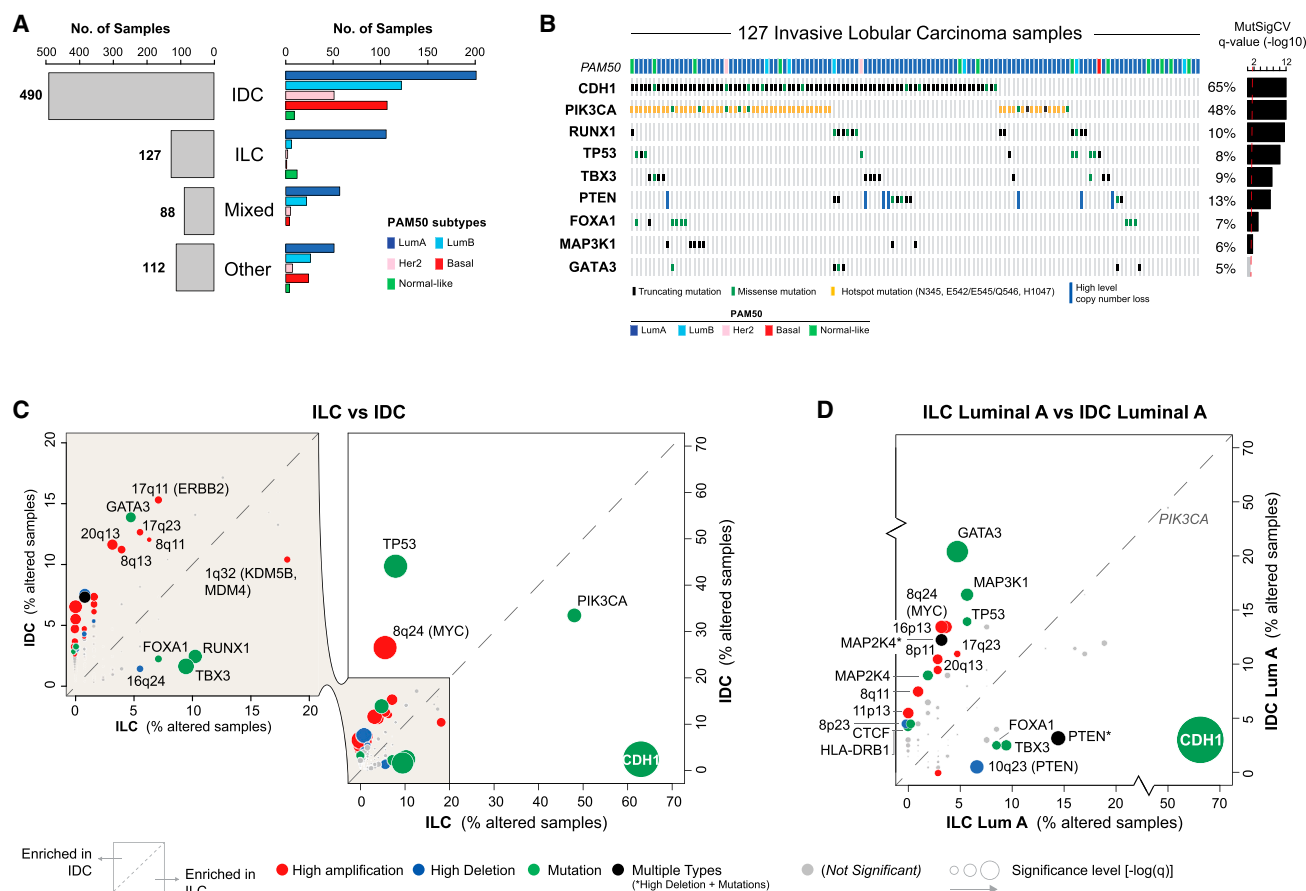


Figure 1. Molecular Determinants of Invasive Lobular Breast Cancer

(A) Histopathological breast cancer subtypes: invasive ductal (IDC), invasive lobular (ILC), mixed ductal/lobular (Mixed), and other-type (Other) carcinoma. PAM50 intrinsic subtypes are not equally distributed across breast cancer subtypes.

(B) Recurrently mutated genes (MutSigCV2) in ILC.

(C) Comparison of the alteration frequency for 153 recurrent genomic alterations in ILC versus IDC.

(D) Comparison of the alteration frequency for 153 recurrent genomic alterations in ILC LumA versus IDC LumA.

Classic ILCs are typically of low histologic grade and low to intermediate mitotic index. They express estrogen and progesterone receptors (ER and PR) and rarely show HER2 protein overexpression or amplification. These features are generally associated with a good prognosis, yet some studies suggest that long-term outcomes of ILC are inferior to stage-matched invasive ductal carcinoma (IDC) (Pestalozzi et al., 2008). Importantly, ILC infiltrative growth pattern complicates both physical exam and mammographic findings and its patterns of metastatic spread often differ from those of IDC (Arpino et al., 2004).

To date, genomic studies of ILC have provided limited insight into the biologic underpinnings of this disease, mostly focusing on mRNA expression and DNA copy-number analysis (McCart Reed et al., 2015). The first TCGA breast cancer study (Cancer Genome Atlas, 2012) reported on 466 breast tumors assayed on six different technology platforms. ILC was represented by only 36 samples, and no lobular-specific features were noted besides mutations and decreased mRNA and protein expression of CDH1. Here, we analyzed nearly twice as many breast tumors

from TCGA (n = 817), including 127 ILC. This study identified multiple genomic alterations that discriminate between ILC and IDC demonstrating at the molecular level that ILC is a distinct breast cancer subtype and providing new insight into ILC tumor biology and therapeutic options.

RESULTS

Genetic Determinants of Invasive Lobular Cancer

A total of 817 breast tumor samples were profiled with five different platforms as previously described (Cancer Genome Atlas Research Network, 2014) and 633 cases were also profiled by reverse-phase protein array (RPPA). A pathology committee reviewed and classified all tumors into 490 IDC, 127 ILC, 88 cases with mixed IDC and ILC features, and 112 with other histologies (Table S1). As expected, lobular tumors were predominantly classified as luminal A (LumA) (Figure 1A) and being typically ER+ tumors characterized by low levels of proliferation markers (Table S1). ER status was clinically determined by

immunohistochemistry on 120 of 127 ILC cases, with 94% ($n = 113$) scoring positively.

Within 127 ILC, we identified 8,173 total coding mutations, integrating information from both DNA and RNA sequencing (Wilkerson et al., 2014). Recurrently mutated genes in ILC were identified by MutSigCV2 (Lawrence et al., 2013) and included many genes previously implicated in breast cancer (Figure 1B, Table 1) (Cancer Genome Atlas, 2012). Similarly, recurrent copy-number alterations in ILC estimated by GISTIC (Mermel et al., 2011) recapitulated known breast cancer gains and losses, in particular those observed in ER+/luminal tumors (Figure S1A). However, the frequency of these alterations (both mutations and copy-number changes) often differed significantly between IDC and ILC.

To investigate these differences, we identified recurrent alterations across all 817 samples and separately in ILC and IDC PAM50 subtypes (luminal A, $n = 201$, luminal B, $n = 122$, HER2-enriched, $n = 51$, and basal-like, $n = 107$). In total, we identified 178 events, including 68 mutated genes, 47 regions of gain, and 63 regions of loss (Table 1 and Table S2). Several of these had different incidence in ILC and in IDC (Figure 1C, Table S3). ILC cases were significantly enriched for *CDH1* mutations (63% in ILC versus 2% in IDC, $q = 3.94E-53$), most of them truncating, and mutations affecting *TBX3* (9% versus 2%, $q = 0.003$), *RUNX1* (10% versus 3%, $q = 0.008$), *PIK3CA* (48% versus 33%, $q = 0.02$), and *FOXA1* (7% versus 2%, $q = 0.08$). By contrast, alterations typically observed in ER-/basal-like tumors were less frequent in ILC, including *TP53* mutations (8% in ILC versus 44% in IDC, $q = 1.9E-14$) and focal amplification of *MYC* (6% versus 27%, $q = 7.42E-7$) and *CCNE1* (0% versus 7%, $q = 0.01$). These results partly reflect genetic differences between ER+/luminal and ER-/basal-like breast cancer, given that ILC tumors were predominantly LumA. Nonetheless, unexpected differences did emerge including a lower incidence of *GATA3* mutations in ILC compared to IDC (5% in ILC versus 13% in IDC, $q = 0.03$) (Figure 1C).

To better identify ILC discriminatory features, we limited our analyses to LumA samples, representing 41% of IDC ($n = 201$) and 83% of ILC ($n = 106$) (Figure 1D). This analysis confirmed a high incidence of *CDH1* ($q = 1.4E-30$), *TBX3* ($q = 0.05$), and *FOXA1* ($q = 0.065$) mutations in ILC, while the frequency of *RUNX1* and *PIK3CA* mutations was no longer significantly different. *GATA3* mutations (5% ILC versus 20% IDC, $q = 0.003$) were the second most discriminant event after *CDH1* mutations, mostly affecting IDC tumors. Interestingly, both *FOXA1* and *GATA3* are key regulators of ER activity (Liu et al., 2014), suggesting IDC and ILC may preferentially rely on different mechanisms to mediate the ER transcriptional program. Finally, homozygous losses of the *PTEN* locus (10q23) were more frequent in ILC ($q = 0.035$) as were *PTEN* mutations (8% versus 3%). Collectively, *PTEN* inactivating alterations were identified in 14% of LumA ILC versus 3% of LumA IDC ($p = 9E-4$), making this the third most discriminant feature between LumA IDC and LumA ILC (Figure 1D).

E-Cadherin Loss in Invasive Lobular Carcinoma

Loss of the epithelial specific cell-cell adhesion molecule E-cadherin (*CDH1*) is the key hallmark of ILC (Dabbs et al., 2013;

Moll et al., 1993). *CDH1* loss is believed to confer the highly discohesive morphology characteristic of this tumor subtype and is often associated with tumor invasion and metastasis in other tumor types, including diffuse gastric cancer (Brinck et al., 2004; Cancer Genome Atlas Research Network, 2014; Richards et al., 1999). Loss-of-function mutations targeting *CDH1* are present in 50%–60% of ILC and are believed to be an early event often observed in matching lobular carcinoma in situ (LCIS) (McCart Reed et al., 2015). *CDH1* mutations typically occur in combination with chromosome 16q loss, where *CDH1* is located, thus inducing complete loss of the protein.

We identified 108 mutations in the coding sequence of *CDH1* in 107/817 patients (63%); 80 of these occurred in ILC cases. These mutations were rather uniformly distributed along the coding sequence, and 83% of them were predicted to be truncating (Figure 2A). *CDH1* mutations almost invariably co-occurred with heterozygous loss of 16q (affecting 89% of ILC cases) and were associated with downregulation of both *CDH1* transcript and protein levels (Figures 2B and S2A). By combining somatic mutations, copy-number losses, and mRNA and low protein expression (the latter when available), we identified E-cadherin alterations in 120/127 (95%) cases with DNA and RNA data, and in all 79 cases with DNA, RNA, and protein data (Figures S2A–SC).

Previous studies reported sporadic cases of multiple cancer types with high DNA methylation levels at the *CDH1* promoter, suggesting epigenetic silencing as an alternative mechanism for downregulation of *CDH1* (Graff et al., 1997; Richards et al., 1999; Sarrió et al., 2003; Zou et al., 2009). We analyzed the DNA methylation levels in breast tumors at CpG sites spanning from upstream of the *CDH1* promoter, across the promoter CpG island, and extending into the first intron (Figure S2D). Despite four of these probes matching DNA positions previously reported as methylated in ILC (Graff et al., 1997; Sarrió et al., 2003; Zou et al., 2009), we did not detect significant DNA hyper-methylation at these probes (Figure S2E), nor in any of the other *CDH1* associated probes analyzed (Figures 2B and S2F and S2G). Moderately increased methylation was observed in a few cases near exon 2; however, DNA methylation at this site correlated with lower tumor purity and increased leukocyte infiltration and indeed it mimicked methylation levels at this CpG site in normal leukocytes (Figures S2F and S2G). Infinium DNA methylation results were validated by whole-genome bisulfite sequencing in five samples (Figure S2H). Altogether, these data confirm that *CDH1* expression was substantially lower in ILC than in IDC and that this expression difference did not appear associated with DNA methylation at the *CDH1* promoter. Our results on 817 invasive breast tumors thus confirmed E-cadherin loss as ILC defining molecular feature but do not support the reported occurrence of *CDH1* epigenetic silencing in invasive breast cancer. The discrepancy with prior literature may be attributable in part to the reliance on highly sensitive, but non-quantitative, methylation-specific PCR assays in past studies (Herman et al., 1996) and will require further investigation.

FOXA1 Mutations in Breast Cancer

FOXA1 is a key ER transcriptional modulator (Carroll et al., 2005; Hurtado et al., 2011) coordinating ER DNA binding within a large

Table 1. Recurrently Mutated Genes in Breast Cancer

Gene	ILC (n = 127)		ILC Luminal A (106)		IDC (490)		IDC Luminal A (201)		IDC Luminal B (122)		IDC Her2-enriched (51)		IDC Basal-like (107)		ALL Breast Cancer (817)	
	n	q	n	q	n	q	n	q	n	q	n	q	n	q	n	q
PIK3CA	61	1.02E-12	54	9.18E-13	164	6.09E-13	93	6.79E-13	43	6.76E-13	19	9.14E-13	7	6.22E-02	282	2.54E-13
RUNX1	13	1.02E-12	9	9.18E-13	13	n.s.	9	1.32E-05	3	n.s.	0	n.s.	1	n.s.	32	2.54E-13
CDH1	80	3.40E-12	68	6.12E-12	10	7.63E-03	7	5.33E-02	2	n.s.	0	n.s.	0	n.s.	107	2.54E-13
TP53	10	8.26E-11	6	2.22E-04	215	6.09E-13	28	6.79E-13	52	6.76E-13	37	9.14E-13	92	1.83E-12	280	2.54E-13
TBX3	12	2.54E-08	10	4.01E-06	8	n.s.	5	n.s.	2	n.s.	0	n.s.	1	n.s.	26	1.11E-08
PTEN	9	8.43E-08	8	8.86E-09	27	5.61E-11	6	5.63E-03	11	9.64E-12	4	3.21E-02	6	4.65E-03	42	2.54E-13
FOXA1	9	5.52E-04	9	6.53E-04	11	n.s.	5	n.s.	3	n.s.	2	n.s.	1	n.s.	30	4.52E-13
MAP3K1	7	2.95E-02	6	7.54E-02	40	4.06E-12	33	1.25E-11	2	n.s.	1	n.s.	4	n.s.	69	2.54E-13
GATA3	6	n.s.	5	n.s.	66	6.09E-13	40	6.79E-13	22	6.76E-13	3	n.s.	0	n.s.	96	2.54E-13
AKT1	3	n.s.	3	n.s.	15	5.61E-11	11	1.25E-11	3	5.91E-02	1	n.s.	0	n.s.	20	2.54E-13
NBL1	3	n.s.	2	n.s.	10	1.08E-10	8	2.04E-12	0	n.s.	1	n.s.	1	n.s.	16	5.24E-11
KMT2C	9	n.s.	8	n.s.	37	1.49E-08	17	2.94E-02	12	n.s.	3	n.s.	5	n.s.	64	4.89E-06
DCTD	0	n.s.	0	n.s.	6	1.02E-05	3	1.54E-02	1	n.s.	1	n.s.	0	n.s.	6	7.61E-04
RB1	0	n.s.	0	n.s.	16	1.46E-04	4	n.s.	7	n.s.	1	n.s.	4	n.s.	18	n.s.
SF3B1	4	n.s.	4	n.s.	12	3.20E-04	6	1.12E-03	4	n.s.	1	n.s.	1	n.s.	16	3.68E-04
CBFB	2	n.s.	2	n.s.	15	1.51E-03	13	5.68E-06	1	n.s.	1	n.s.	0	n.s.	24	8.14E-13
ARHGAP35	1	n.s.	1	n.s.	13	1.62E-03	5	n.s.	5	1.24E-02	0	n.s.	3	n.s.	18	7.74E-03
OR9A2	0	n.s.	0	n.s.	5	1.77E-03	2	n.s.	1	n.s.	1	n.s.	1	n.s.	5	6.47E-03
NCOA3	6	n.s.	4	n.s.	24	1.77E-03	7	n.s.	7	n.s.	4	n.s.	6	n.s.	40	3.25E-07
RBMX	2	n.s.	2	n.s.	10	2.83E-03	3	n.s.	2	n.s.	1	n.s.	4	6.27E-02	12	4.01E-08
MAP2K4	2	n.s.	2	n.s.	24	2.83E-03	18	5.29E-12	5	n.s.	1	n.s.	0	n.s.	30	1.37E-05
TROVE2	0	n.s.	0	n.s.	6	4.51E-03	1	n.s.	2	n.s.	1	n.s.	2	n.s.	8	2.77E-03
NADK	0	n.s.	0	n.s.	4	4.51E-03	0	n.s.	4	5.85E-05	0	n.s.	0	n.s.	6	3.61E-03
CASP8	1	n.s.	1	n.s.	9	6.00E-03	2	n.s.	3	n.s.	0	n.s.	3	n.s.	11	1.81E-03
CTSS	0	n.s.	0	n.s.	5	6.00E-03	1	n.s.	2	n.s.	0	n.s.	2	n.s.	5	8.91E-02
ACTL6B	2	n.s.	1	n.s.	5	7.63E-03	2	n.s.	1	n.s.	0	n.s.	2	n.s.	10	7.33E-05
LGALS1	0	n.s.	0	n.s.	4	9.78E-03	2	n.s.	2	n.s.	0	n.s.	0	n.s.	5	6.34E-03
KRAS	2	n.s.	1	n.s.	4	1.54E-02	3	7.32E-03	0	n.s.	0	n.s.	1	n.s.	7	2.49E-04
KCNN3	2	n.s.	2	n.s.	8	1.81E-02	1	n.s.	2	n.s.	2	n.s.	3	2.45E-02	16	4.53E-02
FBXW7	2	n.s.	2	n.s.	6	2.19E-02	0	n.s.	0	n.s.	0	n.s.	6	8.28E-04	11	n.s.
LRIG2	0	n.s.	0	n.s.	4	3.08E-02	2	n.s.	0	n.s.	1	n.s.	1	n.s.	6	n.s.
PIK3R1	0	n.s.	0	n.s.	9	3.08E-02	2	n.s.	3	n.s.	2	n.s.	2	n.s.	13	1.56E-03
PARP4	3	n.s.	3	n.s.	7	3.08E-02	3	n.s.	4	n.s.	0	n.s.	0	n.s.	12	n.s.
ZNF28	3	n.s.	3	n.s.	7	3.25E-02	1	n.s.	5	n.s.	0	n.s.	1	n.s.	11	1.72E-02
HLA-DRB1	0	n.s.	0	n.s.	13	3.52E-02	9	1.49E-02	2	n.s.	0	n.s.	2	n.s.	16	n.s.
ERBB2	5	n.s.	4	n.s.	7	6.42E-02	3	n.s.	1	n.s.	2	n.s.	1	n.s.	18	3.36E-06
ZMYM3	0	n.s.	0	n.s.	9	8.83E-02	3	n.s.	1	n.s.	1	n.s.	4	n.s.	11	n.s.
RAB42	1	n.s.	1	n.s.	2	n.s.	0	n.s.	0	n.s.	0	n.s.	2	6.27E-02	4	1.82E-03
CTCF	0	n.s.	0	n.s.	12	n.s.	9	7.05E-08	1	n.s.	1	n.s.	1	n.s.	18	1.93E-03
ATAD2	0	n.s.	0	n.s.	9	n.s.	2	n.s.	4	7.32E-02	2	n.s.	1	n.s.	12	n.s.
CDKN1B	3	n.s.	2	n.s.	5	n.s.	4	9.59E-02	1	n.s.	0	n.s.	0	n.s.	11	1.14E-03
GRIA2	0	n.s.	0	n.s.	6	n.s.	5	5.33E-02	0	n.s.	0	n.s.	1	n.s.	6	n.s.
NCOR1	8	n.s.	8	n.s.	23	n.s.	12	4.81E-03	7	n.s.	1	n.s.	3	n.s.	39	3.61E-03
HRNR	4	n.s.	4	n.s.	13	n.s.	3	n.s.	3	n.s.	3	n.s.	4	n.s.	23	7.65E-02
GPRIN2	1	n.s.	1	n.s.	6	n.s.	3	n.s.	1	n.s.	0	n.s.	2	n.s.	11	1.16E-05
PAX2	1	n.s.	0	n.s.	2	n.s.	2	n.s.	0	n.s.	0	n.s.	0	n.s.	4	4.80E-02

(Continued on next page)

Table 1. Continued

Gene	ILC (n = 127)		ILC Luminal A (106)		IDC (490)		IDC Luminal A (201)		IDC Luminal B (122)		IDC Her2-enriched (51)		IDC Basal-like (107)		ALL Breast Cancer (817)	
	n	q	n	q	n	q	n	q	n	q	n	q	n	q	n	q
ACTG1	1	n.s.	1	n.s.	4	n.s.	2	n.s.	0	n.s.	0	n.s.	2	n.s.	8	9.39E-02
AQP12A	0	n.s.	0	n.s.	3	n.s.	1	n.s.	1	n.s.	0	n.s.	1	n.s.	5	2.69E-02
PIK3C3	2	n.s.	2	n.s.	5	n.s.	2	n.s.	0	n.s.	1	n.s.	2	n.s.	11	3.23E-02
MYB	1	n.s.	1	n.s.	7	n.s.	3	n.s.	2	n.s.	0	n.s.	2	n.s.	12	8.91E-02
IRS4	1	n.s.	1	n.s.	6	n.s.	3	n.s.	0	n.s.	0	n.s.	3	n.s.	8	9.38E-02
TBL1XR1	3	n.s.	2	n.s.	3	n.s.	1	n.s.	1	n.s.	1	n.s.	0	n.s.	12	4.71E-04
RPGR	4	n.s.	3	n.s.	11	n.s.	3	n.s.	3	n.s.	2	n.s.	3	n.s.	19	1.26E-03
CCNI	1	n.s.	1	n.s.	2	n.s.	0	n.s.	2	n.s.	0	n.s.	0	n.s.	3	6.93E-02
ARID1A	7	n.s.	5	n.s.	16	n.s.	7	n.s.	4	n.s.	3	n.s.	2	n.s.	33	7.91E-09
CD3EAP	1	n.s.	0	n.s.	2	n.s.	0	n.s.	0	n.s.	0	n.s.	2	n.s.	5	1.29E-02
ADAMTS6	1	n.s.	1	n.s.	3	n.s.	1	n.s.	0	n.s.	0	n.s.	2	n.s.	8	1.81E-03
OR2D2	0	n.s.	0	n.s.	4	n.s.	0	n.s.	3	n.s.	0	n.s.	1	n.s.	5	5.67E-02
TMEM199	0	n.s.	0	n.s.	3	n.s.	0	n.s.	2	n.s.	1	n.s.	0	n.s.	4	3.36E-02
MST1	0	n.s.	0	n.s.	5	n.s.	2	n.s.	2	n.s.	0	n.s.	1	n.s.	7	9.46E-02
RHBG	0	n.s.	0	n.s.	3	n.s.	0	n.s.	0	n.s.	1	n.s.	2	n.s.	4	7.91E-02
ZFP36L1	1	n.s.	1	n.s.	5	n.s.	2	n.s.	2	n.s.	0	n.s.	1	n.s.	8	3.37E-02
TCP11	2	n.s.	0	n.s.	3	n.s.	2	n.s.	0	n.s.	0	n.s.	1	n.s.	6	4.80E-02
CASZ1	4	n.s.	4	n.s.	3	n.s.	0	n.s.	0	n.s.	1	n.s.	2	n.s.	11	2.03E-02
GAL3ST1	1	n.s.	1	n.s.	2	n.s.	0	n.s.	1	n.s.	0	n.s.	1	n.s.	4	7.74E-03
FRMPD2	1	n.s.	1	n.s.	7	n.s.	2	n.s.	4	n.s.	0	n.s.	1	n.s.	9	8.91E-02
GPS2	1	n.s.	1	n.s.	4	n.s.	3	n.s.	0	n.s.	1	n.s.	0	n.s.	8	8.91E-02
ZNF362	0	n.s.	0	n.s.	3	n.s.	3	n.s.	0	n.s.	0	n.s.	0	n.s.	3	8.91E-02

n: number of mutations **q**: MutSigCV2 q value.

protein complex by modifying chromatin accessibility and mediating long-range DNA interactions (Liu et al., 2014). High FOXA1 expression has been previously reported in breast and prostate cancer (Habashy et al., 2008; Sahu et al., 2011) and somatic mutations in the FOXA1 gene have been reported in these tumor types in about 3%–4% of the cases (Barbieri et al., 2012; Cancer Genome Atlas, 2012; Robinson et al., 2015).

Here, we observed a total of 33 FOXA1 mutations in 30/817 (3.7%) tumors (Figure 3A), and the large sample set allowed us to identify regional hotspots in the FOXA1 mutation distribution. Mutations clustered in the fork-head DNA binding (FK) and C terminus transactivation domains (Figure 3A). A similar mutational pattern was observed by combining multiple prostate cancer sequencing studies (Baca et al., 2013; Barbieri et al., 2012; Grasso et al., 2012; Robinson et al., 2015) (Figure S3A), and confirmed by the TCGA prostate cancer project (Robinson et al., 2015). Thus, regional FOXA1 mutation hotspots are selectively altered in a tissue-independent fashion.

Eleven FOXA1 mutations were observed in 9/127 (7%) ILC cases. All FOXA1 mutations in ILC were in the FK domain, whereas mutations in IDC (n = 11) were observed both in FK (n = 6) and other structural elements (n = 5), without a specific preference. The FK domain includes three α helices (H1, H2, H3), three β strands (S1, S2, S3) and two loops, typically referred to as “wings” (W1, W2) (Figures 3B and S3B). FOXA1 mutations

in FK clustered prevalently in the W2 loop. Notable exceptions were recurrently mutated residues I176 (n = 4) and D226 (n = 3). These residues are far from W2 in sequence space and located in different secondary structure elements; however, they are close (within 5 to 10 Å) to residues in W2 in the 3D space (Figures 3C, 3D and S3C). In total, 22 out of 25 FK-mutations in our dataset fall into a restricted 3D space or “mutation structural hotspot” (MSH) (Figure 3D) indicating a selective pressure for targeting protein interactions and functions mediated by this region. Notably, 8/127 ILC cases have FOXA1 mutations within this MSH compared to 4/490 IDC cases (p = 6E-4), further supporting FOXA1 selected mutations as an ILC feature.

FOXA1 DNA binding occurs mostly through helix H3, that recognizes the binding motif and is stabilized by interactions mediated by its “wings” (Cirillo and Zaret, 2007; Gajiwala and Burley, 2000; Kohler and Cirillo, 2010). Only a few residues predicted or experimentally shown to interact with the DNA were mutated (Figures 3B–3E) suggesting that these events are unlikely to affect FOXA1 DNA binding. FOXA1 is a pioneer factor that binds condensed chromatin and triggers DNA demethylation of its binding sites, making them accessible to transcription factors such as ER (Cirillo et al., 2002; Sérandour et al., 2011). FOXA1 activity can therefore be estimated by the methylation status of these sites where occupied FOXA1 DNA binding sites tend to be demethylated. We analyzed DNA methylation levels of the

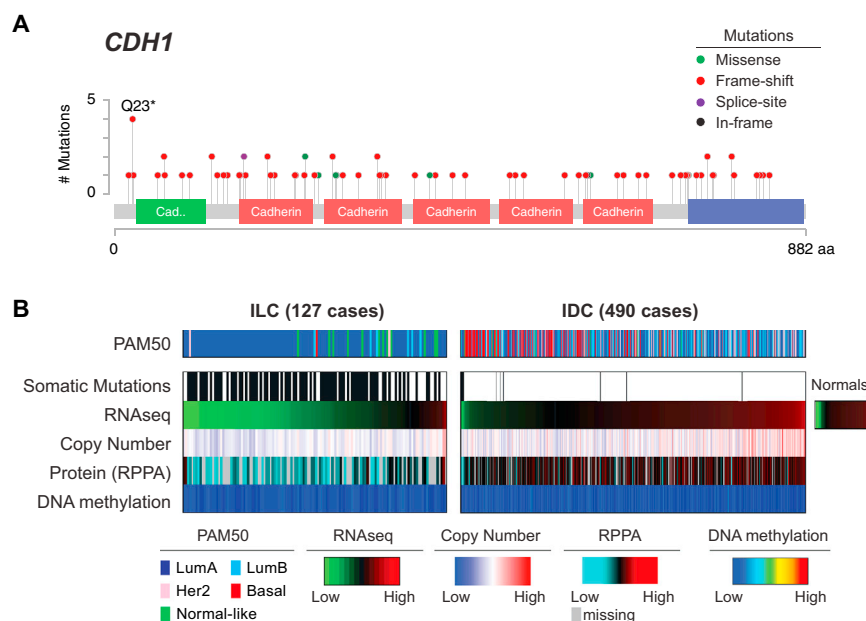


Figure 2. E-Cadherin Loss in ILC

(A) Mutations targeting the CDH1 gene target residues across the whole-sequence and are mostly predicted to be truncating (red). (B) Comparison of E-cadherin status between ILC and IDC reveals frequent hemizygous copy-number losses at the CDH1 locus and downregulation of both mRNA [$\log_2(\text{RSEM})$] and protein levels. See also Figures S2A–S2C. Average DNA methylation level of 6 probes at the CDH1 promoter shows no change in DNA methylation in both ILC and IDC samples. See also Figures S2D–S2I.

3,976 most variable methylation probes mapping to FOXA1 binding sites (Table S3) and methylated in normal samples (Ross-Innes et al., 2012; Wang et al., 2012). DNA methylation of these sites was substantially lower when FOXA1 and ESR1 were highly expressed, while it remained high in FOXA1-negative cases and adjacent normal tissue (Figure 3F). Inverse correlation with FOXA1 expression (Pearson's coefficient $\rho = -0.54$) was specifically observed for DNA methylation at FOXA1 binding sites. Indeed, no correlation was found with methylation at 2,000 most variable probes with the same methylation level in normal samples as FOXA1 binding sites ($\rho = 0.07$) (Figure 3F). These data support the hypothesis that FOXA1 mRNA expression correlates with its activity. FOXA1 mutations were positively associated with its mRNA expression ($p = 0.002$) and maintained a similar anti-correlation with DNA methylation at FOXA1 binding sites (Figure 3F). Finally, by examining mRNA expression of FOXA1 targets, defined as genes with a FOXA1 binding motif in the promoter or matching the genomic loci covered by the 3,976 methylation probes we analyzed (Table S3), no significant differences were identified and only a few genes showed moderate expression changes (Table S4). These data collectively indicate that FOXA1 mutations do not abolish protein function and, in fact, they may activate alternative mechanisms to affect ER transcriptional programs.

Differential expression analyses between FOXA1 mutant and wild-type cases within distinct subsets of samples found consistent upregulation of neuroendocrine secretory proteins SCG1 (CHGB) and SCG2, chemokine-like factor CMTM8, neuroendocrine tumor associated transcription factor NKX2-2 and Kallikrein serine proteases KLK12, KLK13, and KLK14 (Figure S3D and Table S5). While the relatively low number of FOXA1 mutations and breast cancer heterogeneity prevented the identification of strong transcriptional signals associated with FOXA1 mutations, several upregulated targets in FOXA1 mutant cases with

part of them consistently found as significant suggests these lesions might drive novel binding events.

Interestingly, while ILC cases were enriched for FOXA1 mutations, and in particular for those targeting the FK domain, ILC showed significantly fewer GATA3 mutations, another key ER modulator. Mutations in GATA3 were more frequent in LumA IDC (Figure 1D) and mutually

exclusive with FOXA1 events. Moreover, LumA ILC tumors show lower GATA3 mRNA ($p = 0.007$) and protein ($p = 2\text{E-}4$) levels than LumA IDC (Figures S3E and S3F). Taken together, the differential expression patterns and enrichment for hotspot mutations of GATA3 in IDC and of FOXA1 in ILC, suggest a preferential requirement for distinct ER modulators in ILC and IDC.

Akt Signaling Is Strongly Activated in ILC

PTEN inactivation emerged as a discriminant feature between luminal A ILC and luminal A IDC. PTEN genetic alterations across all ILC cases included homozygous deletions (6%) and somatic mutations (7%), and were largely mutually exclusive with PIK3CA mutations (48%) (Figure S4A).

Unbiased differential protein expression analysis (Table S6) based on RPPA data revealed significant lower PTEN protein expression ($p = 4\text{E-}4$) in LumA ILC compared to LumA IDC (Figure 4A). Consistent with PTEN function as a negative regulator of Akt activity (Cantley and Neel, 1999; Song et al., 2012), ILC tumors also showed significantly increased Akt phosphorylation at both S473 ($p = 0.004$) and T308 ($p = 7\text{E-}5$) (Figure 4A). Upstream of the Akt pathway, we found significant upregulation of total EGFR ($p = 1\text{E-}4$) and phospho-EGFR at Y1068 ($p = 0.005$) and Y1173 ($p = 0.007$), as well as phospho-STAT3 at Y705 ($p = 7\text{E-}4$), supporting upregulation of signaling axes converging on Akt activation (Wu et al., 2013). We also identified increased phospho-p27 at T157 ($p = 0.002$), an Akt substrate, and phospho-p70S6 kinase at T389 ($p = 1\text{E-}4$), a direct mTOR target. Notably, ILC phospho-Akt levels were comparable with those typically observed in the more aggressive HER2+ and ER-/basal-like breast tumors (Figure S4B), which have uniformly high levels of PI3K/Akt signaling (Cancer Genome Atlas, 2012). Consistent with these results, we found significant upregulation of a PI3K/Akt pathway-specific protein and phospho-protein expression signature (Akbari et al., 2014) in LumA ILC compared

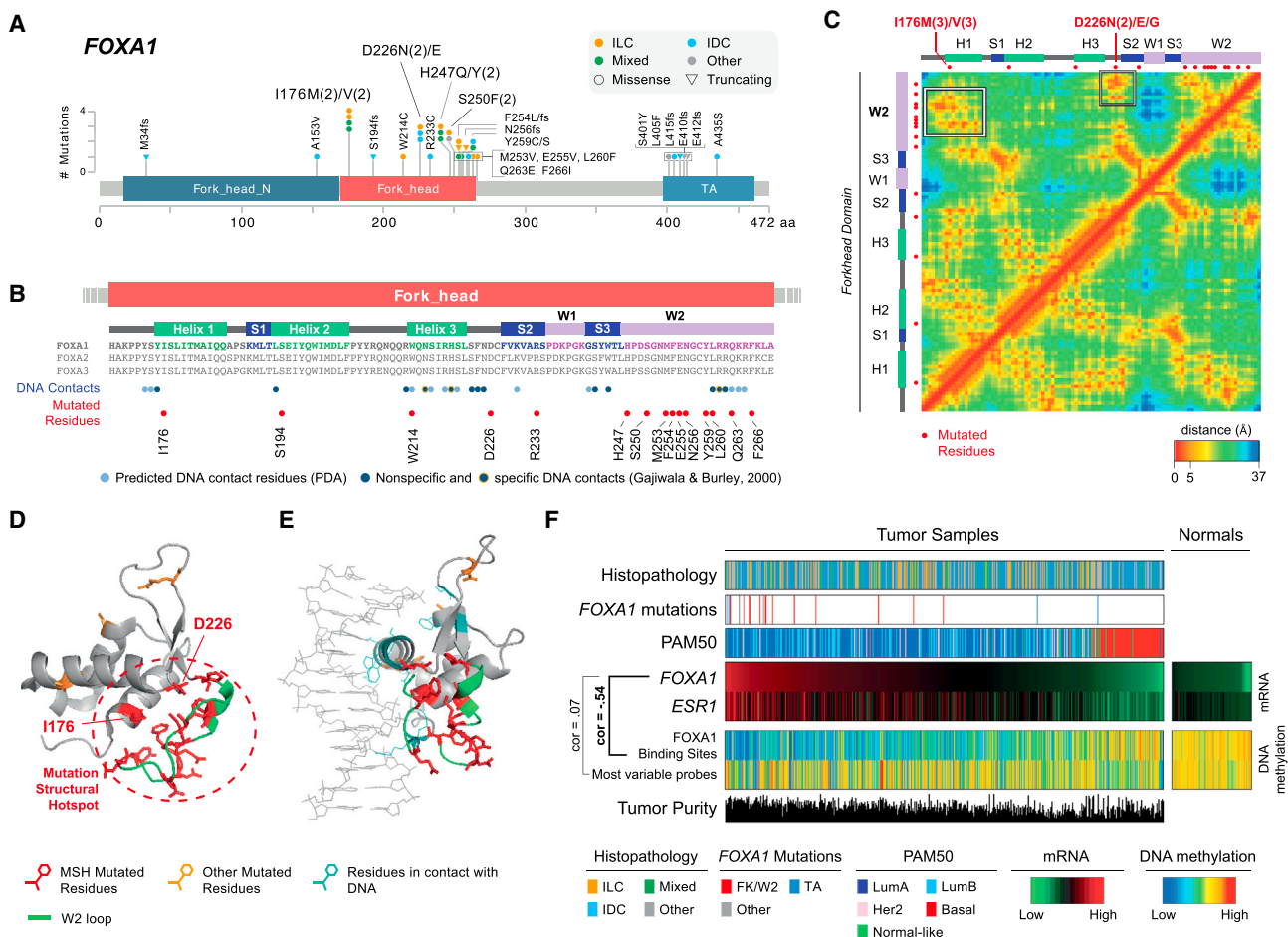


Figure 3. Recurrent FOXA1 Mutations Cluster in the 3D Space and Correlate with High FOXA1 Activity

(A) Recurrent FOXA1 mutations in 817 breast tumors cluster in the Fork-head DNA binding (FK) domain and in the C terminus trans-activation (TA) domain.

(B) Secondary structure elements of the FK domain are not equally mutated. FOXA1 mutations cluster in the W2 loop and rarely target residues interacting with the DNA.

(C) Residue-residue minimum distances for all residues in the FK domain using the 3D structure of FOXA3 FK domain (PDB ID: 1VTN). Frequently mutated residues I176 and D226 are close in the 3D space (but not in the sequence) to the residues in the W2 loop. See also [Figure S3C](#).

(D) 3D structure of the FK domain. Mutations in the W2 loop, in I176, and in D226 form a mutational structural hotspot (MSH).

(E) 3D structure of FK domain bound to the DNA molecule shows mutated residues (red) are not those in contact with the DNA (light blue).

(F) Across all breast cancer subtypes (histopathology and PAM50), FOXA1 mutations are associated with FOXA1 high mRNA expression. FOXA1 mRNA expression is highly correlated with ER mRNA expression [$\log_2(\text{RSEM})$] and anti-correlated with DNA methylation at FOXA1 binding sites consistent with FOXA1 activity. DNA methylation of randomly selected probes was used as control.

to LumA IDC (Figure 4B, Tables S1 and S6). Based on this signature, we found nearly equivalent levels of PI3K/Akt signaling in LumA ILC and basal-like and HER2+ IDC (Figure S4C). Finally, PARADIGM analyses (Vaske et al., 2010) showed increased activation of Akt signaling in LumA ILC relative to LumA IDC (Figure 4D).

PTEN protein loss and increased Akt phosphorylation were observed in association with *PTEN* genetic alterations, as well as in multiple ILC *PTEN* wild-type cases indicating that additional mechanisms contribute to the activation of the pathway. While *PIK3CA* mutations were frequent in LumA ILC tumors, these mutations were not associated with increased levels of phospho-Akt or pathway activity in our dataset. Using MEMO

(Ciriello et al., 2012), we highlighted multiple genetic alterations converging on Akt/mTOR signaling in 45% of the samples (Figure S4D, Table S7). Among these, alterations acting upstream of Akt were identified in 40% of ILC cases and were associated with increased Akt phosphorylation and PI3K/Akt score (Figure 4E), providing an apparent molecular explanation for Akt activation in these samples. Interestingly, these events included *ERBB2* amplification and mutations (Figure 4E), both of which have been identified in relapsed ILC (Ross et al., 2013).

ILC mRNA Subtypes

Using mRNA-seq expression data from LumA ILC samples (n = 106), we identified three ILC subtypes termed *reactive-like*,

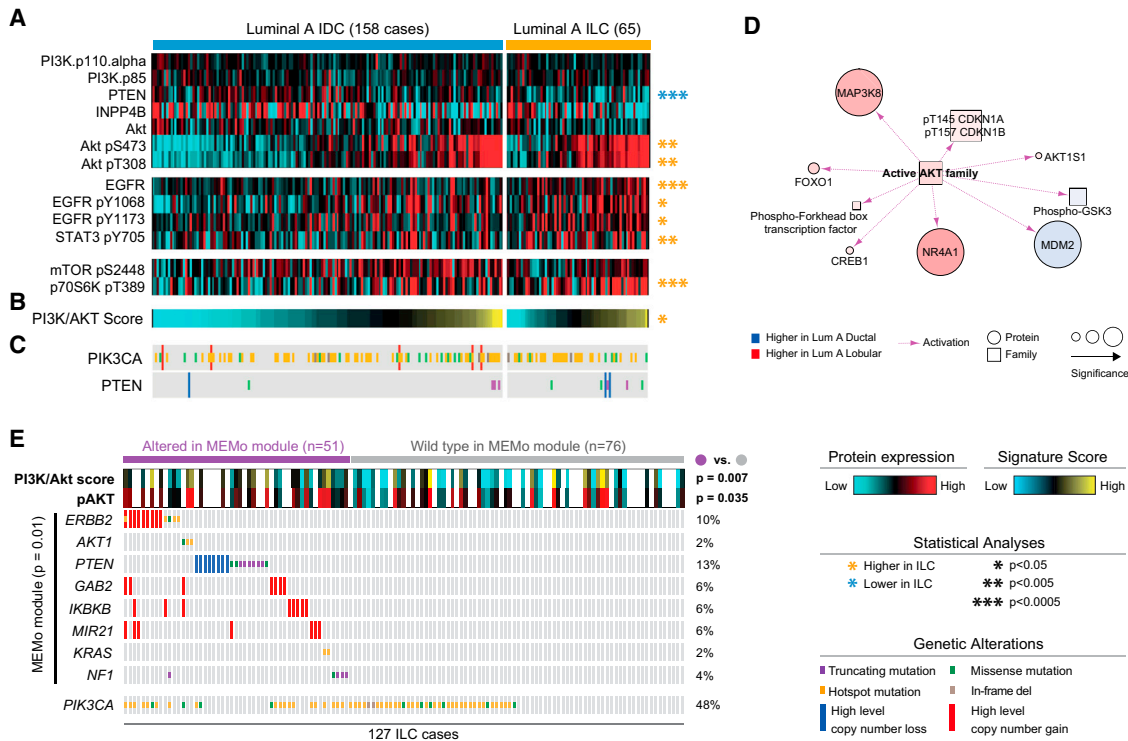


Figure 4. Akt Signaling Is Highest in ILC Tumors

(A) Differential protein and phospho-protein analysis between ILC LumA and IDC LumA reveals significant lower levels of PTEN, and higher levels of Akt, phospho-Akt, EGFR, phospho-EGFR, phospho-STAT3, and phospho-p70S6K in ILC LumA.

(B) A PI3K/Akt protein expression signature is significantly upregulated in ILC tumors. See also Figures S4B–S4C.

(C) Mutation and copy-number alterations in *PIK3CA* and *PTEN*.

(D) PARADIGM identifies increased Akt activity in Luminal A ILC tumors.

(E) MEMo identified multiple mutually exclusive alterations in ILC converging on Akt signaling and associated with increased phospho-Akt and PI3K/Akt protein signature in these tumors. Hotspot are defined as follow: *PIK3CA* E542, E545, Q546, and H1047; *ERBB2* L755, I767, V777; *AKT1* E17; *KRAS* G12.

immune-related, and *proliferative* (Figures S5A–S5I, Table S8). We then used a 3-class ILC subtype classifier (60 genes, Table S13) to score all ILC samples in the TCGA (n = 127) (Figure 5A) and METABRIC (Curtis et al., 2012) datasets (Table S12). Our analyses identified many significant genomic features that distinguished each ILC subtype at the mRNA and protein/phospho-protein level; but no distinguishing somatic mutations or DNA copy-number alterations.

Significant analysis of microarray (SAM) analysis (Tusher et al., 2001) identified 1,277 genes differentially expressed between ILC subtypes (q = 0) (Figure 5A, Table S8). Of these, 1,005 were highly expressed in *reactive-like* tumors, which had lower tumor purity as determined by ABSOLUTE (Carter et al., 2012) (Figures 5A and S5P), and included genes consistent with epithelial and stromal-associated signaling including keratin, kallikrein, and claudin genes as well as the oncogenes *EGFR*, *MET*, *PDGFRA*, and *KIT* (Table S8). The remaining 272 genes were highly expressed in *immune-related* tumors and include modulators of immunogenic signaling such as interleukins (IL), chemokine receptors and ligands, major histocompatibility complex, and tumor necrosis factors, as well as *IDO1* and *IFNG* (Figure 5A and Table S8). Interestingly, immune activity in this subset of tumors appears to be predominantly associated with

macrophage-associated signaling as increased levels of CD68 (p < 0.05), macrophage-associated colony stimulating factor (MacCSF), macrophage-associated TH1 (MacTH1), and T cell receptor (TCR) gene expression signatures (Iglesia et al., 2014) were observed in both the TCGA (Figure 5A) and METABRIC (Figure S5J–S5K) datasets. Finally, *proliferative* tumors were defined by low expression of each of these 1,277 genes (Figure 5A). Intriguingly, in each dataset (Figures 5 and S5K) *proliferative* tumors had higher levels of proliferation relative to *reactive-like* tumors (TCGA: p = 3.3E-09; METABRIC: p = 0.018) and slightly higher or equivalent levels compared to *immune-related* ones (TCGA: p = 0.29; METABRIC: p = 0.008). Regardless of ILC subtype, ILC tumor proliferation was generally lower than all IDC subtypes (Figures S5L–S5M).

With respect to previously reported RPPA-based subtypes (reactive or non-reactive), *reactive-like* ILC largely, but not entirely, comprised tumors classified as reactive (Figure 5B; p < 1E-4), a subgroup characterized by strong microenvironment and/or cancer fibroblast signaling (Cancer Genome Atlas, 2012). Examining protein and phospho-protein expression differences between ILC subtypes identified many significant features (Figure 5B and Table S6). Reactive-like tumors had higher levels of c-Kit (p = 4E-4), consistent with mRNA expression, total

($p = 0.004$) and phosphorylated PKC alpha (S657, $p = 0.002$); beta catenin ($p = 0.012$) and E-cadherin ($p = 0.011$), although both beta-catenin and E-cadherin levels are significantly lower than in all IDC subtypes. Decreased levels were observed instead for p70S6 kinase ($p = 0.017$), Raptor ($p = 0.027$) and eIF4G ($p = 0.024$).

Immune-related tumors had higher levels of immune modulator STAT5 alpha ($p = 0.019$), PI3K/Akt targets phospho-PRAS40 (T246, $p = 0.016$) and mTOR (S2448, $p = 0.019$), and total ($p = 0.004$) and phospho-MEK1 (S217-S221, $p = 0.022$). Consistent with the mRNA proliferation signature, tumors in the *proliferative* subtype have increased expression of cell-cycle proteins cyclin E1 ($p = 0.036$), FoxM1 ($p = 0.019$), PCNA ($p = 0.019$), and prop-sho-Chk1 (S345, $p = 0.038$) as well as DNA repair components Rad50 ($p = 0.007$), Rad51 ($p = 0.007$), XRCC1 ($p = 0.028$), and BRCA2 ($p = 0.038$). Decreased expression was observed for total ($p = 0.014$) and phospho-MAPK (T202-Y204, $p = 0.038$), and phosphorylated MEK1 (S217-S221, $p = 0.019$), PKC alpha (S657, $p = 0.006$), PKC beta (S660, $p = 0.037$), and Src (Y527, $p = 0.026$). Protein pathway signatures (Akbari et al., 2014) recapitulated these findings with *proliferative* tumors having increased levels of the cell cycle ($p = 0.005$) and DNA damage response ($p = 0.014$) signatures and a lower RAS-MAPK signature ($p = 0.031$) score (Tables S1 and S6).

Using an integrative genomics approach, PARADIGM predicted increased activation of the TP53, TP63, TP73 TCF/beta-catenin PKC, and JUN/FOS pathways in *reactive-like* tumors; increased activation of immune-modulators IL12 and IL23, IL12R and IL23R, JAK2 and TYK2 in *immune-related* ILCs (Baay et al., 2011; Duvallet et al., 2011; Strobl et al., 2011), and decreased activation of each of these pathways along with lower levels of MAPK3, RB1, and ERK1 (Figure 5C) in *proliferative* ILC tumors.

Lastly, we determined that *reactive-like* ILC patients had a significantly better disease-specific (DSS) ($p = 0.038$, HR: 0.47) and overall survival (OS) ($p = 0.023$, HR: 0.50) compared to *proliferative* ILC patients in the METABRIC dataset, which has a median follow-up of 7.2 years (compared to the TCGA median follow-up of less than 2 years), (Figure 5D). Consistent with these results, patients with more proliferative lobular tumors (i.e., greater than the median PAM50 proliferation signature score) had worse DSS ($p = 0.025$, HR: 2.0) and a tendency toward worse OS ($p = 0.058$, HR: 0.63) compared to patients with a lower proliferation score (Figures S5N–S5O). No significant differences in DSS or OS were identified between the *immune-related* subgroup and either the *proliferative* or *reactive-like* subgroup. These results are consistent with previous studies reporting that the reactive stromal phenotype is associated with a good prognosis in breast cancer while proliferation is one of the strongest indicators of worse outcome in luminal/ER+ breast cancers (Ciriello et al., 2013b).

Tumors with Mixed ILC and IDC Histology

Histologically, ~3%–6% of breast tumors present both a ductal and a lobular component (Figure 6A). Pathologists currently classify these tumors as *mixed ductal/lobular breast carcinoma* or *invasive ductal cancers with lobular features* (Arps et al., 2013). There are, however, no defined criteria or uniform terminology

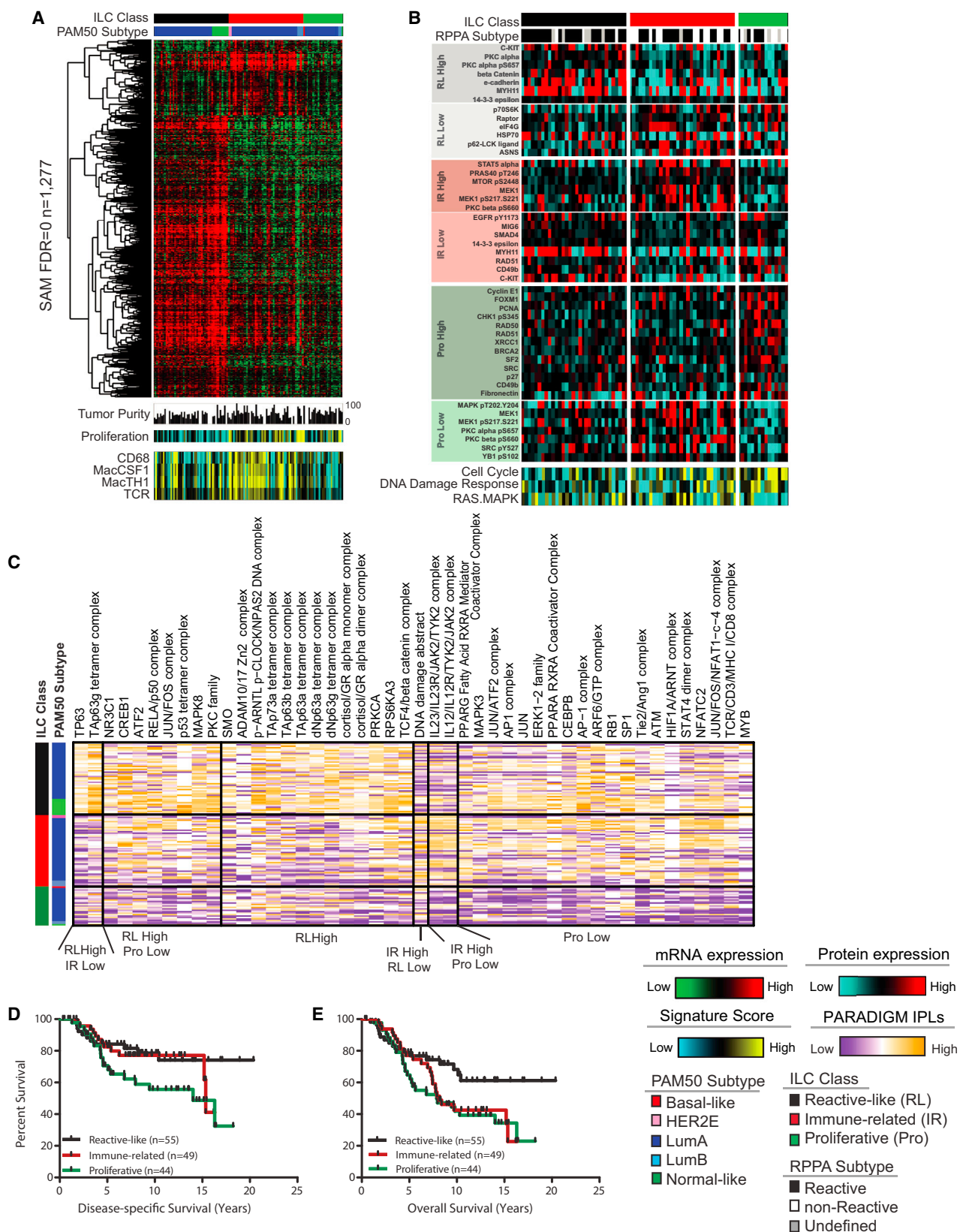
for the classification of *mixed* tumors and as a consequence discordant clinical and molecular features have been reported (Bharat et al., 2009). Molecular profiling has the power to provide quantitative endpoints to compare the genetics of mixed tumors with those of pure ILC and IDC. In our dataset, 88/817 tumors (11%) were classified by as mixed ductal/lobular breast carcinomas. We characterized these *mixed* tumors using multiple computational approaches integrating different data-types thus to determine whether they molecularly resembled IDC (IDC-like), ILC (ILC-like), or neither.

We first analyzed the transcriptional landscape of mixed tumors using the ISOpure algorithm (Quon et al., 2013), which deconvolves the transcriptional signal of each queried tumor to estimate how much of it can be explained by one or more reference populations and how much is unique. Interestingly, mRNA expression profiles of all *mixed* cases could almost completely be explained by either IDC or ILC reference populations, suggesting that these tumors separate into IDC-like and ILC-like cases and do not represent a molecularly distinct subtype (Figure S6A). Based on this analysis, 32/88 *mixed* cases received an ILC-score greater than the IDC-score, and were therefore classified as ILC-like (Figure 6B).

We next evaluated the resemblance of *mixed* tumors to IDC and ILC based on the previously determined selected set of copy-number alterations (CNAs) and mutations (Table S2). *Mixed* tumors were enriched for IDC recurrent CNAs and mutations when compared to ILC, and vice versa (Figure S6B), indicating ILC and IDC genetic alterations were both present in these tumors, either simultaneously or in separate IDC-like and ILC-like subgroups. We then compared each *mixed* tumor to ILC and IDC based on their genomic features, by adapting the OncoSign algorithm (Ciriello et al., 2013a). This approach identified 19 ILC-like mixed samples characterized by ILC genetic features (Figures 6B and S6D). All *CDH1*-mutated *mixed* cases were classified as ILC-like, indicating *CDH1* status as a dominant feature in this analysis. A few *CDH1* wild-type mixed cases were also classified as ILC-like and characterized by ILC-enriched events such as mutations in *RUNX1* (3/4 mutated cases), *TBX3* (2/4), and *FOXA1* (2/6). ILC-enriched alterations did not co-occur with IDC-enriched ones, further indicating that mixed tumors can be categorized into ILC-like or IDC-like subgroups and do not constitute a molecularly distinct subtype.

Finally, we combined 428 CNA, including focal and arm-level alterations, 409 gene expression modules (Fan et al., 2011; Gatz et al., 2014) and somatic mutations for 128 genes mutated in more than 3% of the cases into a single ElasticNet classifier (Zou and Hastie, 2005). This integrated ElasticNet predictor identified 27/88 mixed tumors as ILC-like. These were enriched for the LumA subtype, *CDH1* mutations and loss of E-cadherin mRNA expression (Figures 6B and S6E).

Overall, these approaches were highly concordant (Figures 6B and S6F) with 24/88 cases (18/57 LumA cases) being called ILC-like by at least two approaches, and 64 being called IDC-like (Table S1). ILC-like and IDC-like mixed tumors when compared to pure ILC and IDC, respectively, do not show significant enrichment for specific genomic alterations, being molecularly similar to either one or the other subtype. Our analyses demonstrate that mixed histology tumors overwhelmingly tend



(legend on next page)

to resemble either ILC or IDC as opposed to representing a third distinct group. Moreover, IDC and ILC discriminant molecular features, in particular *CDH1* status, could be used to stratify *mixed* tumors into ILC-like and IDC-like tumor subgroups.

DISCUSSION

In this study we provide the most comprehensive molecular portrait to date of ILC. E-cadherin loss was confirmed the ILC hallmark lesion, and we could identify *CDH1* loss at the DNA, mRNA, and protein level in almost all ILC cases. Moreover, 12/27 *CDH1* mutations in non-ILC cases occurred in *mixed* tumors strongly resembling ILC at the molecular level. Surprisingly, we did not identify DNA hyper-methylation of the *CDH1* promoter in any breast tumor, suggesting that E-cadherin loss is not epigenetically driven. In addition, ILC and IDC differed in the *FOXA1* and *GATA3* mutational spectra, *PTEN* loss, and Akt activation. The lower incidence of *GATA3* mutations in ILC and lower *GATA3* mRNA and protein expression suggest that in Luma ILC tumors there is a preferential occupancy of ER in *FOXA1* bounded sites (Theodorou et al., 2013). Differential ER activity is also observed at the protein level where both total ER ($p = 0.005$) and phospho-ER ($p = 2E-05$) levels are reduced in Luma ILC versus Luma IDC. These findings in the context of recent data suggesting an improved response to the aromatase inhibitor letrozole as compared to tamoxifen in ILC (Metzger et al., 2012; Sikora et al., 2014) warrants further investigation.

The chromatin remodeling factor EP300, also involved in ER modulation, is able to directly acetylate *FOXA1*, and EP300 driven acetylation prevents *FOXA1* DNA binding, but does not affect the protein when already bound (Kohler and Cirillo, 2010). Intriguingly, five acetylation sites have been identified in the wings of the fork-head domain; three of them in W2 (K264, K267, and K270), where most of our newly observed *FOXA1* mutations cluster. These observations lead to the hypothesis that *FOXA1* mutations could alter EP300 dependent acetylation of *FOXA1* without affecting EP300 modulation of ER. While a rigorous evaluation of the role of EP300 in breast cancer and how *FOXA1* mutations interfere with it goes beyond the scope of this study, *FOXA1* mutations, its correlation with *FOXA1* expression and lack of DNA methylation at its binding sites, and exclusivity with *GATA3* mutations support these as events activating *FOXA1* function and, thus, ER transcriptional program.

The PI3K/Akt pathway is among the most altered in cancer providing tumor cells with enhanced growth and survival capa-

bilities. Integrating protein and phospho-protein data with gene expression and pathway activity signatures, we consistently identified increased Akt signaling in ILC versus IDC. Notably, E-cadherin loss has been associated with Akt activation and EGFR overexpression (Lau et al., 2011; Liu et al., 2013). Lack of E-cadherin expression, which characterizes almost all ILC tumors, may thus provide a favorable cellular context for Akt activation. Recently, PI3K and Akt inhibitors entered clinical trials for several cancer types including breast cancer. Here we showed that ILC has on average the highest levels of Akt activation, measured by phospho-Akt and PI3K/Akt signaling among all breast cancer subtypes (comparable to IDC basal-like), making selective inhibition of this pathway in ILC a particularly attractive strategy.

Unbiased characterization of the ILC transcriptome showed a high degree of internal variability giving rise to three main subgroups: *reactive-like*, *immune-related*, and *proliferative*. While additional validation studies will clearly be required, we do observe increased expression of many druggable pathways/targets including increased levels of phospho-mTOR and phospho-MEK1 expression in the *immune-related* subgroup as well as increased SMO and ERK pathway activity in the *reactive-like* subgroup. These results, coupled with difference in clinical outcome, suggest that these subgroups will be important for future studies focused on both the clinical and biological aspects of ILC.

Finally, we showed that *mixed* ILC/IDC tumors could be separated into two major groups based on their molecular resemblance to either ILC (ILC-like) or IDC (IDC-like). The ability to classify cancers with *mixed* phenotypes based on the underlying biology has implications for clinical practice as well as furthering our understanding of the etiology of such lesions. Indeed, ILC carcinomas often metastasize to body sites not colonized by IDCs (e.g., gastrointestinal [GI] tract and peritoneal surfaces). ILC are also typically of low histologic grade and with low to intermediate mitotic index, thus limiting their response to primary chemotherapy (Cristofanilli et al., 2005) and their ability to be detected on PET scans. As such, clinicians must be aware of non-specific symptomatology and favor diagnostic approaches such as anatomical scanning (CT scan) for ILCs. Finally, the identification of ILC enriched molecular features may ultimately lead to the design of ILC-targeted therapies. A more refined classification of mixed cancers as IDC-like or ILC-like will improve our understanding, detection, and follow-up of the disease, and enable a more informed and targeted treatment selection.

Figure 5. ILC Molecular Subtypes

(A) Three molecular subtype of lobular breast cancer were identified based on differential gene expression and show unique patterns highly expressed genes ($n = 1277$, SAM FDR = 0, upper panel), minor difference in tumor purity measured by ABSOLUTE, and differences in gene expression signatures measuring proliferation, CD68, Macrophage-associated CSF1, Macrophage-associated TH1, and T Cell Receptor Signaling (lower panel). Proliferation is highest in the *proliferative* (Pro) and *immune-related* (IR) subgroups; macrophage associated signaling is highest in *immune-related* tumors.

(B) Differences in protein expression profiles as determined by RPPA analysis. The *reactive-like* (RL) subgroup shows a significant association ($p < 1E-4$, Fisher's Exact test) with the RPPA-defined reactive subgroup of breast cancer. Differences in subgroup-specific patterns of protein expression ($p < 0.05$, t test) for individual proteins (upper panel) as well as for protein expression signatures (lower panel) were identified. The proliferative subgroup shows higher expression of the cell-cycle and DNA damage response pathways and lower levels of Ras-MAPK signaling ($p < 0.05$).

(C) Subgroup-associated signaling features identified by PARADIGM.

(D and E) *Reactive-like* ($n = 55$) tumors have significantly better (D) disease specific ($p = 0.038$, HR: 0.47, log-rank) and (E) overall survival ($p = 0.022$, HR = 0.50) compared to *proliferative* ($n = 44$) tumors in the METABRIC cohort.

See also Figure S5.

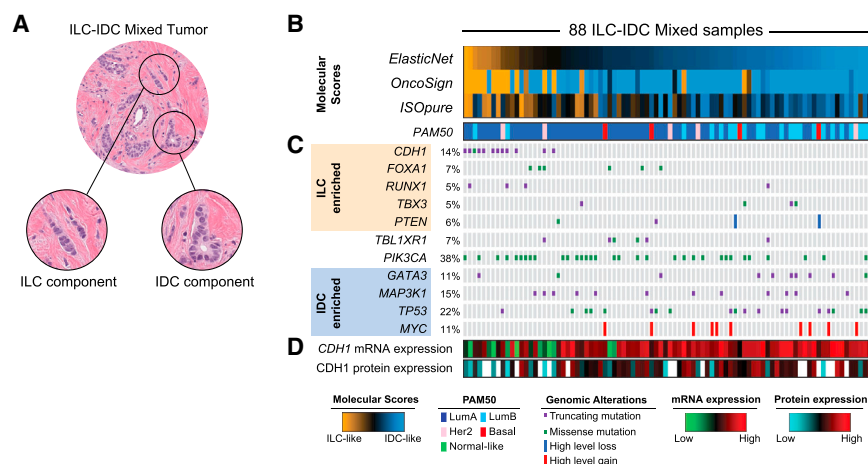


Figure 6. Molecular Classification of Mixed Ductal/Lobular Carcinoma

(A) Mixed ductal/lobular tumors present at the same time both a lobular and ductal component. (B) We used three algorithmic approaches (ElasticNet, OncoSign, and ISOpure) to evaluate the resemblance of mixed tumors to either ILC (ILC-like) or IDC (IDC-like) based on molecular features. ILC-IDC scores are shown for all three approached at the top. See also Figures S6B–S6D. (C) Genetic alterations enriched in ILC tumors are frequently found in ILC-like mixed cases (in particular CDH1 mutations), whereas those enriched in IDC are more frequent in IDC-like mixed cases. (D) ILC-like mixed-cases are characterized by both low E-cadherin mRNA and protein level. See also Figure S6E.

This multi-platform study identified numerous molecular features discriminating between breast ILC and IDC, demonstrating different pathways underlying their pathogenesis, defining new ILC subtypes with different clinical outcomes, and pointing to previously unrecognized therapeutic possibilities. Importantly, we provided here a curated and integrated dataset for 817 breast tumors, including the largest collection to date of comprehensively profiled ILC. To facilitate the exploration of this dataset, we created a public web-service (http://cbio.mskcc.org/cancergenomics/tcga/brca_tcga) organizing all analyses and data used in this manuscript. We believe this resource will serve as a reference for many to further advance our understanding of human breast cancers.

EXPERIMENTAL PROCEDURES

Tumor and matched normal specimen were collected as previously described (Cancer Genome Atlas, 2012). In total 817 primary tumor samples were assayed by whole-exome DNA sequencing, RNA sequencing, miRNA sequencing, SNP arrays, and DNA methylation arrays. A subset of 633 samples was assayed by reverse phase protein array (RPPA). Histological subtypes have been determined based on consensus by a pathology committee. Intrinsic breast cancer subtyping was performed on all 817 cases, using the PAM50 classifier (Parker et al., 2009). Data generation and processing were performed as previously described (Cancer Genome Atlas Research Network, 2014).

Enrichment analyses for selected events were performed using Fisher's exact tests and a binary representation of copy-number alterations and mutations (1 is altered, 0 is wild-type).

DNA methylation of the CDH1 promoter was assessed at probes within a window 1,500 bp upstream and downstream CDH1 transcription start site using both HM27 and HM450 data. Whole-genome bisulfite sequencing was performed to characterize DNA methylation levels at 157 CpGs.

Distances between FOXA1 mutations have been determined from the tertiary structure of FOXA3 fork-head domain (PDB ID: 1VTN). Predicted DNA interactions were derived by WebPDA (<http://bioinfozen.unc.edu/webpda>). Differential expression analyses on RNA-seq data were performed using the limmavoom package (Law et al., 2014).

Replication Based Normalized (RBN) RPPA data containing expression levels for 187 protein and phosphorylated proteins for 633 samples were used for protein differential expression analysis. Differential pathway activity was assessed by t test.

ILC subtypes were determined using Consensus Cluster Plus Analysis (Wilkerson and Hayes, 2010) based on the 1,000 most differentially expressed genes and a classifier was built using ClaNC (Dabney, 2006).

Detailed description of each analysis presented in this study can be found within the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and nine tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.09.033>.

AUTHOR CONTRIBUTIONS

G.C., M.L.G. and C.M.P. coordinated overall study design and analyses. A.H.B., S.C.L., G.M.K.T., R.E.F., L.C.C., K.H.A., Y.-Y.C., K.J., and N.B.J. coordinated pathology analyses. M.D.W., M.McL., and C.K. coordinated DNA sequencing analyses. M.L.G., K.A.H. and C.M.P. coordinated RNA-seq analyses. S.K.R., H.S., and P.W.L. coordinated DNA methylation analyses. H.Z. and A.D.C. coordinated copy-number analyses. R.B. and G.R. coordinated miRNA-seq analyses. G.B.M. coordinated RPPA analyses. A.P., C.Y., S.H., R.F. were involved in bioinformatics analyses. C.B. and C.S. supervised bioinformatics analyses. S.O. and T.A.K. coordinated clinical contributions. G.C. developed web-resource for breast cancer data. G.C., M.L.G. and C.M.P. wrote the manuscript, which all authors reviewed.

CONSORTIA

Rehan Akbani, J. Todd Auman, Miruna Balasundaram, Saianand Balu, Thomas Barr, Andrew Beck, Christopher Benz, Stephen Benz, Mario Berrios, Rameen Beroukhi, Tom Bodenheimer, Lori Boice, Moiz S. Bootwalla, Jay Bowen, Reanne Bowlby, Denise Brooks, Andrew D. Cherniack, Lynda Chin, Juok Cho, Sudha Chudamani, Giovanni Ciriello, Tanja Davidsen, John A. Demchok, Jennifer B. Dennison, Li Ding, Ina Felau, Martin L. Ferguson, Scott Frazer, Stacey B. Gabriel, JianJiong Gao, Julie M. Gastier-Foster, Michael L. Gatz, Nils Gehlenborg, Mark Gerken, Gad Getz, William J. Gibson, D. Neil Hayes, David I. Heiman, Katherine A. Hoadley, Andrea Holbrook, Robert A. Holt, Alan P. Hoyle, Hai Hu, Mei Huang, Carolyn M. Hutter, E. Shelley Hwang, Stuart R. Jefferys, Steven J.M. Jones, Zhenlin Ju, Jaegil Kim, Phillip H. Lai, Peter W. Laird, Michael S. Lawrence, Kristen M. Leraas, Tara M. Lichtenberg, Pei Lin, Shiyun Ling, Jia Liu, Wenbin Liu, Laxmi Lolla, Yiling Lu, Yussanne Ma, Dennis T. Maglinte, Elaine Mardis, Jeffrey Marks, Marco A. Marra, Cynthia McAllister, Michael McLellan, Shaowu Meng, Matthew Meyerson, Gordon B. Mills, Richard A. Moore, Lisle E. Mose, Andrew J. Mungall, Bradley A. Murray, Rashi Naresh, Michael S. Noble, Steffi Oesterreich, Olufunmilayo Olopade, Joel S. Parker, Charles M. Perou, Todd Pihl, Gordon Saksena, Steven E. Schumacher, Kenna R. Mills Shaw, Nilsa C. Ramirez, W. Kimryn Rathmell, Suhan K. Rhie, Jeffrey Roach, A. Gordon Robertson, Gordon Saksena, Chris Sander, Jacqueline E. Schein, Nikolaus Schultz, Hui Shen, Margi Sheth, Yan Shi,

Juliann Shih, Carl Simon Shelley, Craig Shriver, Janae V. Simons, Heidi J. Sofia, Matthew G. Soloway, Carrie Sougne, Charlie Sun, Roy Tarnuzzer, Daniel G. Tiezzi, David J. Van Den Berg, Doug Voet, Yunhu Wan, Zhining Wang, John N. Weinstein, Daniel J. Weisenberger, Matthew D. Wilkerson, Richard Wilson, Lisa Wise, Maciej Wiznerowicz, Junyuan Wu, Ye Wu, Liming Yang, Christina Yau, Travis I. Zack, Jean C. Zenklusen, Hailei Zhang, Jiashan Zhang, Erik Zmuda.

ACKNOWLEDGMENTS

We wish to thank the many members of the TCGA Network, including the Tissue Source Sites, and patients, whom contributed samples to this study. This study was supported by funds from the TCGA Project (U24-CA143848), the NCI Breast SPORE program grant P50-CA58223-09A1, and the Breast Cancer Research Foundation. G.C. is supported by the Gabriella Giorgi-Cavaglieri Foundation, M.L.G. is supported by the National Cancer Institute of the US NIH award number K99-CA166228, A.H.B. is supported by National Library of Medicine of the NIH under Award Number K22LM011931, and A.P. is supported by Mildred-Scheel Postdoctoral Research Fellowship of the Deutsche Krebshilfe e.V. (No. 111354). C.M.P. is an equity stock holder, and Board of Director Member, of BioClassifier. C.M.P. is also listed an inventor on patent applications on the Breast PAM50 assay. A.D.C. receives research funding from Bayer AG.

Received: June 10, 2015

Revised: August 4, 2015

Accepted: September 10, 2015

Published: October 8, 2015

REFERENCES

- Akbani, R., Ng, P.K., Werner, H.M., Shahmoradgoli, M., Zhang, F., Ju, Z., Liu, W., Yang, J.Y., Yoshihara, K., Li, J., et al. (2014). A pan-cancer proteomic perspective on The Cancer Genome Atlas. *Nat. Commun.* 5, 3887.
- Arpino, G., Bardou, V.J., Clark, G.M., and Elledge, R.M. (2004). Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. *Breast Cancer Res.* 6, R149–R156.
- Arps, D.P., Healy, P., Zhao, L., Kleer, C.G., and Pang, J.C. (2013). Invasive ductal carcinoma with lobular features: a comparison study to invasive ductal and invasive lobular carcinomas of the breast. *Breast Cancer Res. Treat.* 138, 719–726.
- Baay, M., Brouwer, A., Pauwels, P., Peeters, M., and Lardon, F. (2011). Tumor cells and tumor-associated macrophages: secreted proteins as potential targets for therapy. *Clin. Dev. Immunol.* 2011, 565187.
- Baca, S.C., Prandi, D., Lawrence, M.S., Mosquera, J.M., Romanel, A., Drier, Y., Park, K., Kitabayashi, N., MacDonald, T.Y., Ghandi, M., et al. (2013). Punctuated evolution of prostate cancer genomes. *Cell* 153, 666–677.
- Barbieri, C.E., Baca, S.C., Lawrence, M.S., Demichelis, F., Blattner, M., Theurillat, J.P., White, T.A., Stojanov, P., Van Allen, E., Stransky, N., et al. (2012). Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat. Genet.* 44, 685–689.
- Bharat, A., Gao, F., and Margenthaler, J.A. (2009). Tumor characteristics and patient outcomes are similar between invasive lobular and mixed invasive ductal/lobular breast cancers but differ from pure invasive ductal breast cancers. *Am. J. Surg.* 198, 516–519.
- Brinck, U., Jacobs, S., Neuss, M., Tory, K., Rath, W., Kulle, B., and Füzesi, L. (2004). Diffuse growth pattern affects E-cadherin expression in invasive breast cancer. *Anticancer Res.* 24, 2237–2242.
- Cancer Genome Atlas, N.; Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70.
- Cancer Genome Atlas Research Network (2014). Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513, 202–209.
- Cantley, L.C., and Neel, B.G. (1999). New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway. *Proc. Natl. Acad. Sci. USA* 96, 4240–4245.
- Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., et al. (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* 122, 33–43.
- Carter, S.L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P.W., Onofrio, R.C., Winckler, W., Weir, B.A., et al. (2012). Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* 30, 413–421.
- Ciriello, G., Cerami, E., Sander, C., and Schultz, N. (2012). Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.* 22, 398–406.
- Ciriello, G., Miller, M.L., Aksoy, B.A., Senbabaoglu, Y., Schultz, N., and Sander, C. (2013a). Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* 45, 1127–1133.
- Ciriello, G., Sinha, R., Hoadley, K.A., Jacobsen, A.S., Reva, B., Perou, C.M., Sander, C., and Schultz, N. (2013b). The molecular diversity of Luminal A breast tumors. *Breast Cancer Res. Treat.* 141, 409–420.
- Cirillo, L.A., and Zaret, K.S. (2007). Specific interactions of the wing domains of FOXA1 transcription factor with DNA. *J. Mol. Biol.* 366, 720–724.
- Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jamik, M., and Zaret, K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9, 279–289.
- Cristofanilli, M., Gonzalez-Angulo, A., Sneige, N., Kau, S.W., Broglio, K., Theriault, R.L., Valero, V., Buzdar, A.U., Kuerer, H., Buchholz, T.A., and Hortobagyi, G.N. (2005). Invasive lobular carcinoma classic type: response to primary chemotherapy and survival outcomes. *J. Clin. Oncol.* 23, 41–48.
- Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al.; METABRIC Group (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486, 346–352.
- Dabbs, D.J., Schnitt, S.J., Geyer, F.C., Weigelt, B., Baehner, F.L., Decker, T., Eusebi, V., Fox, S.B., Ichiara, S., Lakhani, S.R., et al. (2013). Lobular neoplasia of the breast revisited with emphasis on the role of E-cadherin immunohistochemistry. *Am. J. Surg. Pathol.* 37, e1–e11.
- Dabney, A.R. (2006). ClaNc: point-and-click software for classifying microarrays to nearest centroids. *Bioinformatics* 22, 122–123.
- Duvallet, E., Semerano, L., Assier, E., Falgarone, G., and Boissier, M.C. (2011). Interleukin-23: a key cytokine in inflammatory diseases. *Ann. Med.* 43, 503–511.
- Fan, C., Prat, A., Parker, J.S., Liu, Y., Carey, L.A., Troester, M.A., and Perou, C.M. (2011). Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. *BMC Med. Genomics* 4, 3.
- Foote, F.W., Jr., and Stewart, F.W. (1946). A histologic classification of carcinoma of the breast. *Surgery* 19, 74–99.
- Gajiwala, K.S., and Burley, S.K. (2000). Winged helix proteins. *Curr. Opin. Struct. Biol.* 10, 110–116.
- Gatza, M.L., Silva, G.O., Parker, J.S., Fan, C., and Perou, C.M. (2014). An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast cancer. *Nat. Genet.* 46, 1051–1059.
- Graff, J.R., Herman, J.G., Myöhänen, S., Baylin, S.B., and Vertino, P.M. (1997). Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. *J. Biol. Chem.* 272, 22322–22329.
- Grasso, C.S., Wu, Y.M., Robinson, D.R., Cao, X., Dhanasekaran, S.M., Khan, A.P., Quist, M.J., Jing, X., Lonigro, R.J., Brenner, J.C., et al. (2012). The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 487, 239–243.
- Habashy, H.O., Powe, D.G., Rakha, E.A., Ball, G., Paish, C., Gee, J., Nicholson, R.I., and Ellis, I.O. (2008). Forkhead-box A1 (FOXA1) expression in breast cancer and its prognostic significance. *Eur. J. Cancer* 44, 1541–1551.
- Herman, J.G., Graff, J.R., Myöhänen, S., Nelkin, B.D., and Baylin, S.B. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. USA* 93, 9821–9826.

- Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., and Carroll, J.S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* 43, 27–33.
- Iglesia, M.D., Vincent, B.G., Parker, J.S., Hoadley, K.A., Carey, L.A., Perou, C.M., and Serody, J.S. (2014). Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. *Clin. Cancer Res.* 20, 3818–3829.
- Kohler, S., and Cirillo, L.A. (2010). Stable chromatin binding prevents FoxA acetylation, preserving FoxA chromatin remodeling. *J. Biol. Chem.* 285, 464–472.
- Lau, M.T., Klausen, C., and Leung, P.C. (2011). E-cadherin inhibits tumor cell growth by suppressing PI3K/Akt signaling via β -catenin-Egr1-mediated PTEN expression. *Oncogene* 30, 2753–2766.
- Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29.
- Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218.
- Liu, X., Su, L., and Liu, X. (2013). Loss of CDH1 up-regulates epidermal growth factor receptor via phosphorylation of YBX1 in non-small cell lung cancer cells. *FEBS Lett.* 587, 3995–4000.
- Liu, Z., Merkurjev, D., Yang, F., Li, W., Oh, S., Friedman, M.J., Song, X., Zhang, F., Ma, Q., Ohgi, K.A., et al. (2014). Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell* 159, 358–373.
- McCart Reed, A.E., Kutasovic, J.R., Lakhani, S.R., and Simpson, P.T. (2015). Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. *Breast Cancer Res.* 17, 12.
- Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12, R41.
- Metzger, O., Giobbie-Hurder, A., Mallon, E., Viale, G., Winer, E., Thürlimann, B., Gelber, R.D., Colleoni, M., Ejlersen, B., Bonnefoi, H., et al. (2012). Abstract S1-1: Relative effectiveness of letrozole compared with tamoxifen for patients with lobular carcinoma in the BIG 1-98 trial. In Thirty-Fifth CTRC-AACR San Antonio Breast Cancer Symposium (San Antonio, TX).
- Moll, R., Mitze, M., Frixen, U.H., and Birchmeier, W. (1993). Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. *Am. J. Pathol.* 143, 1731–1742.
- Morrough, M., Andrade, V.P., Giri, D., Sakr, R.A., Paik, W., Qin, L.X., Arroyo, C.D., Brogi, E., Morrow, M., and King, T.A. (2012). Cadherin-catenin complex dissociation in lobular neoplasia of the breast. *Breast Cancer Res. Treat.* 132, 641–652.
- Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27, 1160–1167.
- Pestalozzi, B.C., Zahrieh, D., Mallon, E., Gusterson, B.A., Price, K.N., Gelber, R.D., Holmberg, S.B., Lindtner, J., Snyder, R., Thürlimann, B., et al.; International Breast Cancer Study Group (2008). Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: combined results of 15 International Breast Cancer Study Group clinical trials. *J. Clin. Oncol.* 26, 3006–3014.
- Quon, G., Haider, S., Deshwar, A.G., Cui, A., Boutros, P.C., and Morris, Q. (2013). Computational purification of individual tumor gene expression profiles leads to significant improvements in prognostic prediction. *Genome Med.* 5, 29.
- Richards, F.M., McKee, S.A., Rajpar, M.H., Cole, T.R., Evans, D.G., Jankowski, J.A., McKeown, C., Sanders, D.S., and Maher, E.R. (1999). Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. *Hum. Mol. Genet.* 8, 607–610.
- Robinson, D., Van Allen, E.M., Wu, Y.M., Schultz, N., Lonigro, R.J., Mosquera, J.M., Montgomery, B., Taplin, M.E., Pritchard, C.C., Attard, G., et al. (2015). Integrative clinical genomics of advanced prostate cancer. *Cell* 161, 1215–1228.
- Ross, J.S., Wang, K., Sheehan, C.E., Boguniewicz, A.B., Otto, G., Downing, S.R., Sun, J., He, J., Curran, J.A., Ali, S., et al. (2013). Relapsed classic E-cadherin (CDH1)-mutated invasive lobular breast cancer shows a high frequency of HER2 (ERBB2) gene mutations. *Clin. Cancer Res.* 19, 2668–2676.
- Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481, 389–393.
- Sahu, B., Laakso, M., Ovaska, K., Mirtti, T., Lundin, J., Rannikko, A., Sankila, A., Turunen, J.P., Lundin, M., Konsti, J., et al. (2011). Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. *EMBO J.* 30, 3962–3976.
- Sarrió, D., Moreno-Bueno, G., Hardisson, D., Sánchez-Estévez, C., Guo, M., Herman, J.G., Gamallo, C., Esteller, M., and Palacios, J. (2003). Epigenetic and genetic alterations of APC and CDH1 genes in lobular breast cancer: relationships with abnormal E-cadherin and catenin expression and microsatellite instability. *Int. J. Cancer* 106, 208–215.
- Sérandour, A.A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Métié, R., et al. (2011). Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res.* 21, 555–565.
- Sikora, M.J., Cooper, K.L., Bahreini, A., Luthra, S., Wang, G., Chandran, U.R., Davidson, N.E., Dabbs, D.J., Welm, A.L., and Oesterreich, S. (2014). Invasive lobular carcinoma cell lines are characterized by unique estrogen-mediated gene expression patterns and altered tamoxifen response. *Cancer Res.* 74, 1463–1474.
- Song, M.S., Salmena, L., and Pandolfi, P.P. (2012). The functions and regulation of the PTEN tumour suppressor. *Nat. Rev. Mol. Cell Biol.* 13, 283–296.
- Strobl, B., Stoiber, D., Sexl, V., and Mueller, M. (2011). Tyrosine kinase 2 (TYK2) in cytokine signalling and host immunity. *Front. Biosci. (Landmark Ed.)* 16, 3214–3232.
- Theodorou, V., Stark, R., Menon, S., and Carroll, J.S. (2013). GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* 23, 12–22.
- Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* 98, 5116–5121.
- Vaske, C.J., Benz, S.C., Sanborn, J.Z., Earl, D., Szeto, C., Zhu, J., Haussler, D., and Stuart, J.M. (2010). Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 26, i237–i245.
- Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* 22, 1798–1812.
- Wilkerson, M.D., and Hayes, D.N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573.
- Wilkerson, M.D., Cabanski, C.R., Sun, W., Hoadley, K.A., Walter, V., Mose, L.E., Troester, M.A., Hammerman, P.S., Parker, J.S., Perou, C.M., and Hayes, D.N. (2014). Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. *Nucleic Acids Res.* 42, e107.
- Wu, K., Chang, Q., Lu, Y., Qiu, P., Chen, B., Thakur, C., Sun, J., Li, L., Kowluru, A., and Chen, F. (2013). Gefitinib resistance resulted from STAT3-mediated Akt activation in lung cancer cells. *Oncotarget* 4, 2430–2438.
- Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc., B* 67, 301–320.
- Zou, D., Yoon, H.S., Perez, D., Weeks, R.J., Guilford, P., and Humar, B. (2009). Epigenetic silencing in non-neoplastic epithelia identifies E-cadherin (CDH1) as a target for chemoprevention of lobular neoplasia. *J. Pathol.* 218, 265–272.

Ligand-Dependent Enhancer Activation Regulated by Topoisomerase-I Activity

Janusz Puc, Piotr Kozbial, Wenbo Li, Yuliang Tan, Zhijie Liu, Tom Suter, Kenneth A. Ohgi, Jie Zhang, Aneel K. Aggarwal, and Michael G. Rosenfeld*

*Correspondence: mrosenfeld@ucsd.edu

<http://dx.doi.org/10.1016/j.cell.2015.08.044>

(Cell 160, 367–380; January 29, 2015)

Our paper studied the role of the DNA nickase TOP1 in enhancer activation and transcription of eRNA in androgen receptor (AR)-regulated transcriptional programs in prostate cancer cells. The central findings of the paper were that TOP1 is recruited to AR-bound active enhancers, along with elements of the DNA damage repair machinery, and that DNA nicking activity of TOP1 is required for robust eRNA synthesis and enhancer activation.

It has come to our attention that, in Figures 5 and S5A, showing the kinetics of recruitment of several members of the DNA damage repair machinery to the KLK3, KLK2, TMPRSS2, and NDRG1 enhancer regions, the presented panels showing the results of the qPCR of the Ku80 and p-S1983-ATM ChIPs were identical. We retrieved the original files used to prepare the figures in the paper and were able to verify that the values depicted in the p-S1983-ATM graphs were incorrect. We reconstructed all steps of the preparation of the published figures and concluded that the errors originated when data were compiled using GraphPad software and we mistakenly copied the Ku80 values into p-S1983-ATM plots. Using the raw data from the original experiments, we were able to plot the correct numbers and we now provide corrected versions of [Figures 5](#) and [S5](#) below. The figures have also been corrected online. The correct data support the original interpretation that pATM is maximally recruited to enhancers ~60 min after dihydrotestosterone-induced binding of androgen receptor, and the conclusions of the paper remain unchanged. We apologize for not identifying these errors before publication and for any inconvenience that they may have caused.

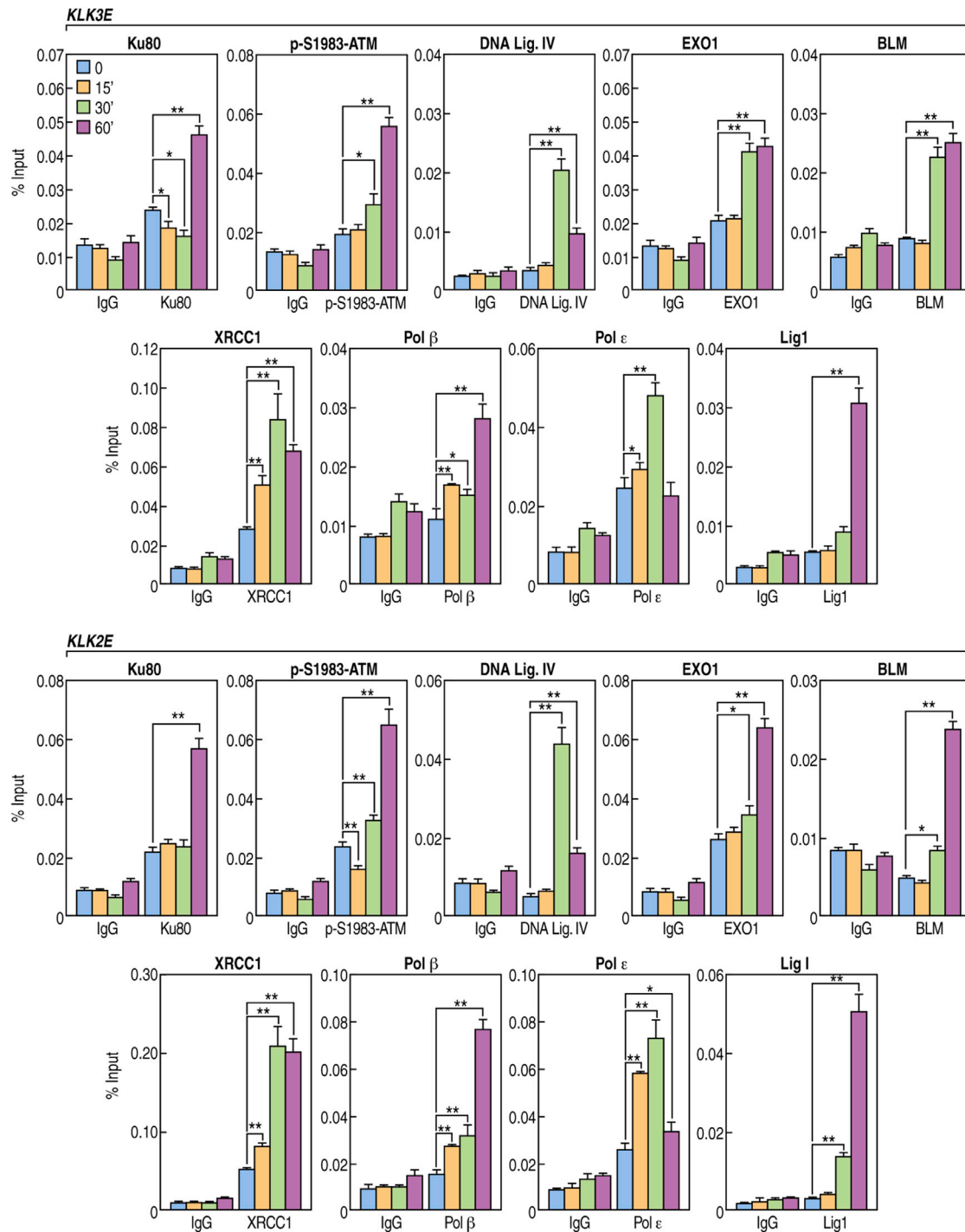


Figure 5. Canonical DNA Damage/Repair Machinery Components Recruit to AR-Regulated Enhancers

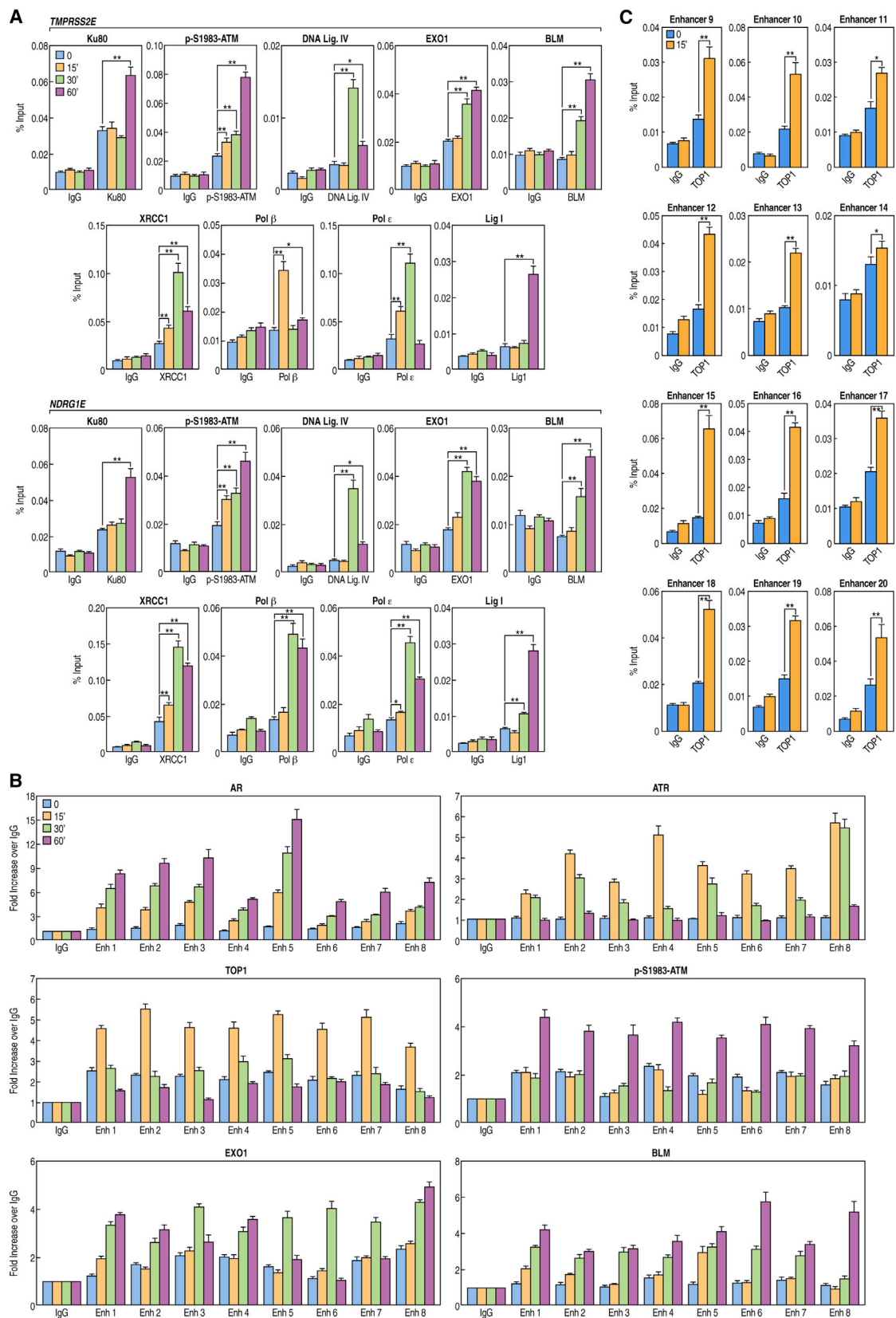


Figure S5. Kinetic ChIP Studies Performed on Additional AR Enhancers, Related to Figure 5

Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses

Roi Avraham, Nathan Haseley, Douglas Brown, Cristina Penaranda, Humberto B. Jijon, John J. Trombetta, Rahul Satija, Alex K. Shalek, Ramnik J. Xavier, Aviv Regev, and Deborah T. Hung*

*Correspondence: hung@molbio.mgh.harvard.edu
<http://dx.doi.org/10.1016/j.cell.2015.09.044>

(Cell 162, 1309–1321; September 10, 2015)

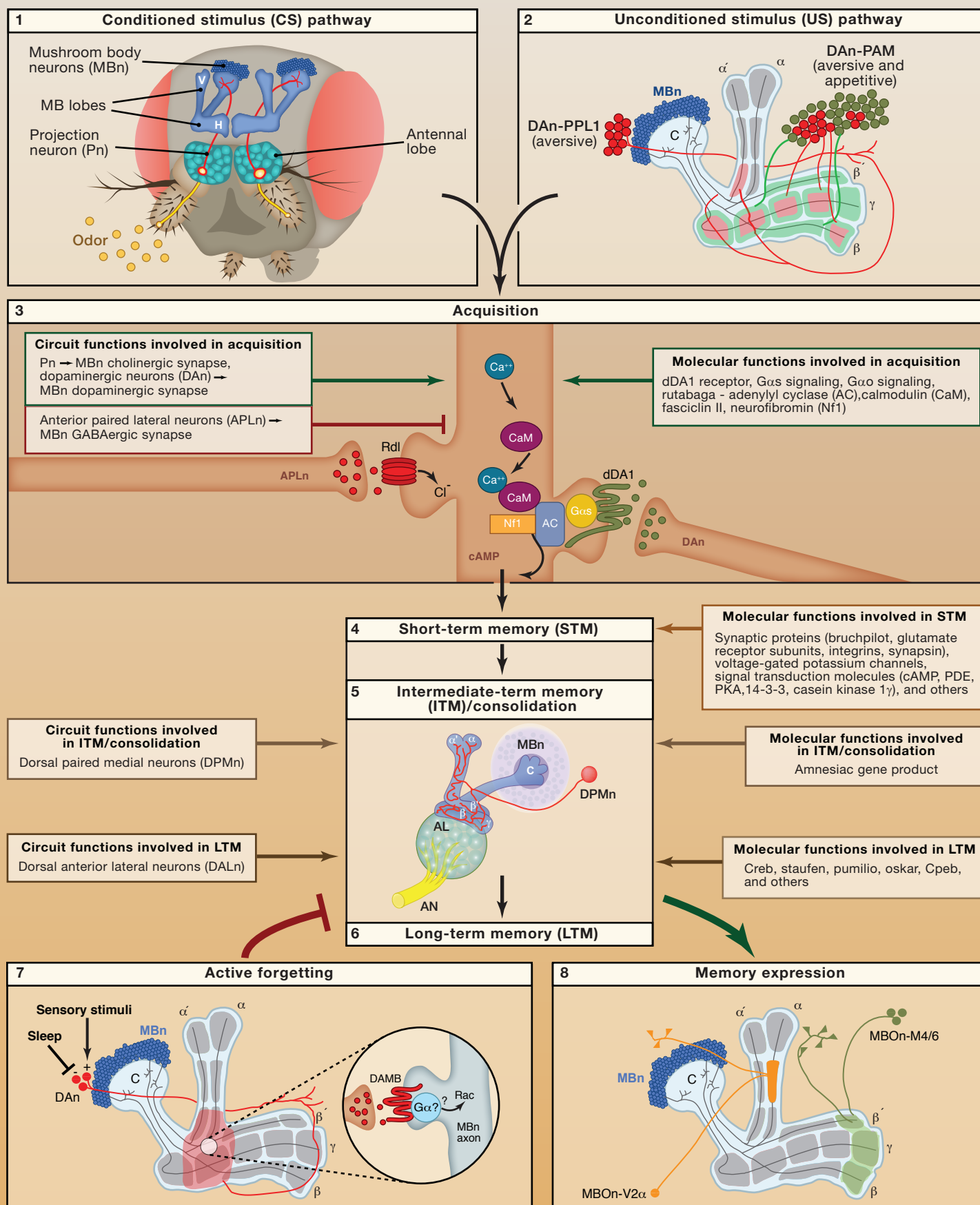
In the above article, the authors inadvertently included an incorrect GEO accession number, GSE66528. The article has been corrected online to reflect the correct accession number, GEO: GSE65528.

SnapShot: Olfactory Classical Conditioning of *Drosophila*

Cell

Ronald L. Davis

Department of Neuroscience, The Scripps Research Institute Florida, Jupiter, FL 33458, USA



SnapShot: Olfactory Classical Conditioning of *Drosophila*

Cell

Ronald L. Davis

Department of Neuroscience, The Scripps Research Institute Florida, Jupiter, FL 33458, USA

Classical conditioning with odor cues occurs when a specific odor conditioned stimulus (CS) is paired with an aversive unconditioned stimulus (US) such as electric shock or an appetitive US such as a food reward (Tomchik and Davis, 2013). This conditioning produces avoidance behavior to odors associated with an aversive US and approach behavior to those associated with an appetitive US. The memory components from such conditioning include acquisition; short-, intermediate- (consolidation), and long-term memory; active forgetting; and memory expression. Many different types of neurons within the olfactory nervous system—brain structures primarily involved in the processing of odorant cues—contribute to olfactory memory formation.

1. Conditioned Stimulus Pathway: The sensation and perception of specific odors (yellow circles) occurs from the activation of olfactory receptor neurons (ORN) in the antenna. Odor identity is established by the activation of odor-specific sets of projection neurons (Pn, red) located in the antennal lobe (AL, turquoise), the equivalent of the olfactory bulb in vertebrates (Masse et al., 2009). The Pn transmit odor identity to synapses (red branches) with the mushroom body neurons (MBn, blue circles) in the dorsal posterior brain. The MBn fall into three major classes, the α/β , α'/β' , and γ MBn (see 2). Their axons project through neuropil known as the horizontal (H) and vertical (V) MB lobes (blue), named according to the class of axons they contain. MBn are thus third order in the olfactory pathway (Tomchik and Davis, 2013).

2. Unconditioned Stimulus Pathway: Accumulated evidence indicates that dopaminergic neurons (DAn) are necessary and sufficient to serve as the US signal (Waddell, 2013). Two clusters of DAn (PPL1 and PAM) serve this purpose and innervate segments of the horizontal and vertical lobes. Those DAn providing aversive and appetitive US are colored red and green, respectively, along with the MB lobe regions that they innervate.

3. Acquisition: Acquisition occurs, in part, through the temporal coincidence of the odor CS and the aversive or appetitive US. MBn are intimately involved in acquisition. Information about the odor CS conveyed by Pn is transmitted to the MBn by the activation of MBn-expressed ACh receptors. This leads to an increase in intracellular calcium concentration. The US information is transmitted from DAn to MBn using the dopamine receptor, dDA1. Increased calcium, along with the activation of the dDA1 receptor on the MBn, produces a synergistic increase in cAMP concentration. The mobilization of the cAMP signaling system is essential for acquisition. Many different molecules are required for acquisition, including the relevant receptors, calmodulin-activated adenylyl cyclase, neurofibromin, and others. GABAergic input to receptors expressed by MBn from the APLn constrains acquisition (Tomchik and Davis, 2013).

4. Short-Term Memory (STM): Short-term memory is brought about, in part, through plastic changes in the MBn and other cell types due to the transient increase in cAMP level. Genetics and molecular biology have identified many different molecules required for STM, including various synaptic proteins, ion channels, cAMP signaling molecules, and others. A memory trace corresponding to STM, observed as an increased calcium response to the CS+ odor after aversive conditioning, is observed in the axons of the α'/β' MBn (Tomchik and Davis, 2013; Davis, 2011).

5. Intermediate-Term Memory (ITM)/Consolidation: Consolidation into protein-synthesis-dependent LTM occurs across the first few hours after conditioning, making it difficult to distinguish effects of mutations on a specific temporal form of memory—ITM—or the process of consolidation. ITM/consolidation requires the reciprocal interactions between dorsal paired medial neurons (DPMn) and the MBn. The *amnesiac* gene product, expressed by the DPMn, is required for the expression of protein-synthesis-dependent LTM. A memory trace corresponding to ITM is observed in the processes of the DPMn innervating the vertical MB lobes (Tomchik and Davis, 2013; Davis, 2011).

6. Long-Term Memory (LTM): Protein-synthesis-dependent LTM requires the actions of the transcription factor, Creb, in the MBn along with genetic functions for local translation (staufen, pumilio, oskar, Cpeb) and a host of other genetic functions specifically involved in LTM. In addition to the MBn, the dorsal anterior lateral neurons (DALn, not shown) are required for LTM. A memory trace corresponding to LTM is observed in the vertical axons of the α/β MBn (Dubnau et al., 2003; Davis, 2011).

7. Active Forgetting: One mechanism for active forgetting involves release of dopamine (DA) from a small group of DAn (red) clustered near the MB calyx (C). These DAn innervate the MBn axons in an area joining the vertical and horizontal lobes. DA activates a receptor named DAMB that initiates the forgetting of memory traces stored in the MBn. This DA signal is suppressed by the state of sleep and is activated by sensory stimuli (Berry et al., 2015). The intracellular signaling for this receptor is at present unclear but may involve the downstream activation of the small G protein, Rac (Shuai et al., 2010).

8. Memory Expression: Synaptic output from a small group of glutamatergic neurons (MBO_n-M4/6, green) that project dendrites into the distal tips of the horizontal lobes is required for the expression of both aversive ITM and appetitive LTM. These neurons project their axons to a terminal field residing between the vertical and horizontal lobes. A memory trace, observed as an increased calcium response to the CS+ odor after aversive conditioning and a decreased response to the CS+ odor after appetitive conditioning, is observed in the dendrites of these neurons (Owald et al., 2015). Synaptic output from a small group of cholinergic neurons (MBO_n-V2 α , orange) that project dendrites into a defined segment of the vertical lobe is also required for the expression of aversive ITM and LTM (Séjourné et al., 2011). These neurons project their axons to lateral regions of the brain. A memory trace, observed as a decreased calcium response to the CS+ odor after aversive conditioning, is observed in the dendrites of these neurons. The relative activity of these and other MBO_n with dendritic arbors in other segments of the MB lobes is thought to capture and integrate the activity of MBn to help drive avoidance or approach behavior (Aso et al., 2014).

ABBREVIATIONS

CS, conditioned stimulus; US, unconditioned stimulus; ORn, olfactory receptor neuron; Pn, projection neuron; MBn, mushroom body neuron; DA, dopamine; DAn, dopamine neuron; ACh, acetylcholine; APLn, anterior paired lateral neuron; STM, short-term memory; ITM, intermediate-term memory; AN, antennal nerve; LTM, long-term memory; DPMn, dorsal paired medial neuron; DALn, dorsal anterior lateral neuron; C, mushroom body calyx; MBO_n, mushroom body output neuron.

ACKNOWLEDGMENTS

Dr. J. Berry produced some of the original artwork for a prior publication. The artwork has been modified and reproduced with permission from Berry and Davis (Progress in Brain Research 208, 293–304) and Davis (Neuron 70, 8–19). I thank S. Tomchik and J. Berry for comments. Research was supported by R37NS19904-32 and 5R01NS052351-09.

REFERENCES

- Aso, Y., Sitaraman, D., Ichinose, T., Kaun, K.R., Vogt, K., Belliard-Guérin, G., Plaçais, P.Y., Robie, A.A., Yamagata, N., Schnaitmann, C., et al. (2014). eLife 3, e04580. 10.7554/eLife.04580.
- Berry, J.A., Cervantes-Sandoval, I., Chakraborty, M., and Davis, R.L. (2015). Cell 167, 1656–1667.
- Davis, R.L. (2011). Neuron 70, 8–19.
- Dubnau, J., Chiang, A.S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., et al. (2003). Curr. Biol. 13, 286–296.
- Masse, N.Y., Turner, G.C., and Jefferis, G.S. (2009). Curr. Biol. 19, R700–R713.
- Owald, D., Felsenberg, J., Talbot, C.B., Das, G., Perisse, E., Huetteroth, W., and Waddell, S. (2015). Neuron 86, 417–427.
- Séjourné, J., Plaçais, P.Y., Aso, Y., Siwanowicz, I., Trannoy, S., Thoma, V., Tedjakumala, S.R., Rubin, G.M., Tchénio, P., Ito, K., et al. (2011). Nat. Neurosci. 14, 903–910.
- Shuai, Y., Lu, B., Hu, Y., Wang, L., Sun, K., and Zhong, Y. (2010). Cell 140, 579–589.
- Tomchik, S.M., and Davis, R.L. (2013). Learn. Mem. 22, 357–377.
- Waddell, S. (2013). Curr. Opin. Neurobiol. 23, 324–329.